TURING 图灵新知

宇宙之书

从托勒密、爱因斯坦到多重宇宙

The Book of Universes

〔英〕约翰·D. 巴罗/著　李剑龙/译



人民邮电出版社
POSTS & TELECOM PRESS

TURING 图灵新知

宇宙之书——从托勒密、爱因斯坦到多重宇宙

The Book of Universes

<英> 约翰·D.巴罗/著 李剑龙/译

人民邮电出版社

北京

图书在版编目（**CIP**）数据
宇宙之书：从托勒密、爱因斯坦到多重宇宙/（英）巴罗著；李剑龙译.--北京：人民邮电出版社，2013.4
（图灵新知）
书名原文：The book of universes
ISBN　978-7-115-30972-3
Ⅰ.①宇…　Ⅱ.①巴…②李…　Ⅲ.①宇宙—普及读物　Ⅳ.①P159-49
中国版本图书馆CIP数据核字（2013）第049799号

内容提要

头上的星空，自古以来就激发着人类无尽的好奇和想象。为了解释我们所见的 宇宙，从亚里士多德、托勒密、哥白尼、牛顿、康德到爱因斯坦，再到现代学者对 爱因斯坦宇宙学方程组的艰难求解，人们提出了各式各样的理论，描述了种种面貌 不一的宇宙。而在最新的多重宇宙论中，可能的宇宙不只限于理论，而是真实的存在， 数量更数不胜数，其中的自然法则可能各不相同。身为知名的宇宙学家和科普作家， 作者将在这本宇宙之书中带领我们回顾历史，追踪现代天文学的前沿进展，纵览“奇 异得超乎我们想象”的万千宇宙。

图灵新知

宇宙之书——从托勒密、爱因斯坦到多重宇宙

◆著　[英] 约翰·D. 巴罗
译　李剑龙
责任编辑　楼伟珊
◆人民邮电出版社出版发行　　北京市崇文区夕照寺街14号
邮编　100061　　电子邮件　315@ptpress.com.cn
网址　http://www.ptpress.com.cn
北京　　印刷
开本：880×1230　1/32
◆印张：11.25
字数：291字　　2013年4月第1版
印数：1-4000册　　2013年4月北京第1次印刷

著作权合同登记号　图字：01-2011-2963号

ISBN　978-7-115-30972-3

定价：39.00元

读者服务热线：（**010**）**51095186**转**604**　印装质量热线：（**010**）**67129223**

反盗版热线：（**010**）**67171154**

版权声明

献给蒂莉
许多宇宙将会呈现在你面前

推荐序一

在广义相对论和现代宇宙学的研究方面，英国有很强的学术传统，特别是剑桥大学，一直是国际上宇宙学研究的一个主要中心。从爱丁顿开始，到稳恒态宇宙学的主将霍伊尔，再到以发现黑洞辐射而著名的霍金，英国有许多一流的宇宙学家。而且英国宇宙学家还有一个优良传统，就是他们当中很多人（包括上面提到的几位）都非常喜爱且擅长撰写科普作品。这些作品不仅内容新颖，介绍了宇宙学研究的前沿成果，而且语言风趣幽默，影响了世界各地的许多青少年。有不少宇宙学家，最早就是因为读了这些书而对宇宙学发生兴趣的。

约翰· D.巴罗教授也是一位多产的英国宇宙学家。截至 2012年底，他已经撰写了四百八十多篇论文、二十多本书，还为报刊、杂志等写了很多评论。巴罗教授在宇宙学模型及其观测检验方面完成了大量的研究工作，特别是在检验可变的引力常数方面。我自己的博士毕业论文研究也在这一领域。1999年我从哥伦比亚大学毕业时，巴罗刚刚就任剑桥大学应用数学与理论物理学系教授，我获得了一个到他那里做博士后的机会。不过，由于那时我刚刚开始和一位女孩谈恋爱，因此我选择了留在美国而没有去他那里——补偿我的是，这位女孩后来成了我妻子。直到2010年夏天，我才有机会再去剑桥访问，并在那里见到了他。

巴罗教授有几本科普书已被译成中文，在剑桥的访问结束时他赠给了我几本，使我得以愉快地消磨了回国飞机上的一段时光。这几本书并不像教科书那样系统地介绍某个理论或某种知识，而是有点像一位博学的朋友和你在咖啡馆聊天那样，东拉西扯却又让人兴味盎然。对于一个典型的理科男，也就是那些对数学、物理、天文、哲学、历史之类内容感兴趣的人来说，它们无疑是非常饶有兴味的消遣读物。

本书在某些方面的风格与那几本书也有些相似，里面有不少作者颇费了一番工夫收集甚至考证出来的关于学者们的各种奇闻轶事和八卦，还有许多本来和主题风马牛不相及但被引用时又让人有画龙点睛之感的幽默引语。不过，就内容而言，本书却相当系统。正如它的标题所示，它介绍了从古老的托勒密时代开始直到最近，研究者们为了理解宇宙而构想出的各种各样稀奇古怪的宇宙理论。现在对宇宙学感兴趣的一般读者们不难从许多科普书籍中找到对标准的大爆炸宇宙学理论的介绍，但本书所介绍的许多宇宙学理论，比如说在有些维度上膨胀而在另一些维度上收缩的卡斯纳宇宙、允许时间旅行的哥德尔旋转宇宙、九种类型的均匀而又各向异性的比安基宇宙等，则很少在科普书籍上露面，甚至对

于一些专业的宇宙学研究者而言，如果不做理论研究的话，往往也只是在广义相对论教科书的脚注或者习题里才偶尔与之邂逅。既然这些模型并非主流，甚至有些已被淘汰，那么为什么要在一本面向公众的书中加以介绍呢？科学是在不同学说的竞争中发展的，现代被学者们普遍接受的标准宇宙模型，也是在经过观测和实验检验，击败了众多对手之后才脱颖而出的。但由于科普书籍通常不得不做的过于简化的处理，不少喜爱思索的读者在读完介绍大爆炸宇宙学的书籍后往往并不能完全满足而会产生这样的疑问：你们宇宙学家是有一些观测的证据，但难道对这些观测就不容许有其他的解释吗？读完本书，你可能对科学发展的历史会有一种不同的认识：其实，学者们一直都在创建不同的宇宙学理论；许多学说尽管在竞争中落败了，但对它们的研究，同样丰富和加深了我们对宇宙的理解。

本书也介绍了多重宇宙和人择原理这些仍处在激烈争议中的“后现代”宇宙学说。对这类学说，许多——也许是大部分——从事专业宇宙学研究的人都不太喜欢，认为它们充满了玄学味道，其论证随时可能陷入逻辑漏洞，也无法被观测和实验所证伪——而可证伪性不正是真正科学理论的标志吗？但无可否认的是，不管你喜不喜欢，从逻辑出发，多重宇宙和人择原理是可能的。对这些问题的思索和研究，虽然很难取得那种被学术界一致认可的发现，但对于加深我们对宇宙的理解，却仍是很有必要且也是非常有趣的。不过，关于这些问题的论证往往是颇为抽象难懂的，而且往往还夹杂着一般人所不熟悉的暴胀宇宙学和弦论中一些复杂的技术问题和术语。本书的介绍是比较清晰而平易的，抛开了这些技术上的细节而更多地指向问题的本质。当然，这个“平易”是相对而言的，它还是要求读者具有相当的知识基础。比如，下面这段话是我感到关于人择原理最简明的解释：“重点不在于宇宙拥有特定属性的概率，而在于宇宙中可能存在观测者（“生命”）的同时，也拥有这种特定属性的条件概率。”一下子就点透了人择原理最基本的思想。当然，你必须很熟悉或者至少是了解“条件概率”的概念才会对这段话有这种“通透”的感觉。

总之，巴罗教授的这本著作是一本颇为有趣、让人大开眼界的宇宙学科普书籍。无论是初次接触宇宙学的人，还是对之有一定了解甚至从事专业研究的人，都可以从阅读此书中得到一些乐趣。我愿向所有对宇宙学感兴趣的读者们推荐本书。

陈学雷

中国科学院国家天文台

推荐序二

约翰·D.巴罗比我年长十岁，作为科普作家，起步比我早了近二十年。他的第一部科普著作发表于1983年，眼下这本我为之写序的书是2011年出版的。他是高产的宇宙学家，发表了四百多篇论文，也是高产的科普作家，出版了近二十本科普著作，同时还是一位业余剧作家，至少有一部剧作上演过。

所有这些只能说明他精力过人，也许并不说明任何其他事。我知道，他与物理学家弗兰克·梯普勒（Frank J. Tipler）合著了一本关于宇宙人择原理的书，这本书就人择原理的历史清本溯源，涉及了神学与智能设计，同时也是一本天体物理的原创学术著作。我想，巴罗不论在宇宙学研究上还是科普写作上都是一位有独到创见和视角的人。

我没有读过巴罗的其他科普著作，眼下这本书也许并不能代表他的一贯写作风格。这本书主要为读者介绍多重宇宙论，一个目前在部分宇宙学家中流行但还没有观测证据的理论。但作者还是从我们头顶的星空谈起，从古老的亚里士多德宇宙论、托勒密系统谈起，再到哥白尼的革命，一直谈到爱因斯坦宇宙学，甚至大爆炸前一刻的奇点问题、暴胀宇宙论、膜世界，最后才是多重宇宙以及该学说带来的问题。

我将此书推荐给一切对世界和宇宙起源，包括我们人类自身起源抱有好奇心的读者，但我不敢担保每位读者都能真正理解或领悟书中的多数概念和名词后面的真正含义。也许，阅读有时只扩展我们的眼界，增进我们对这个难以理解同时可以理解世界的好奇。当然，阅读此书还有一个功效，就是给我们带来庞杂的知识，就像本书作者用引语显示的那样，你不知道他的水到底有多深——因为他不仅引用科学家，还引用作家、报纸甚至时尚。

如果有物理系本科训练甚至研究生训练，这本书绝对是一本不可多得的、能够帮助我们梳理宇宙学这个领域各式各样理论的饶有趣味的书。作者显示的庞杂的跨界知识可以用来作为阅读的佐餐——但不必在文学高度寄以期待。

李淼
中国科学院理论物理研究所

前言

时下，关于宇宙的书可谓洋洋大观。这又是一本关于宇宙的书，但我们的故事将围绕着一个不同寻常却未曾被重视的事实展开，那就是阿尔伯特·爱因斯坦（1879～1955）已经向我们展示了怎样描述可能的宇宙——从整体上描述宇宙。在此之前，关于我们所生活的宇宙拥有什么样的结构，人们已经讨论了几千年。为了描述和解释它，人们建立了各式各样奇怪的图景，只是这些图景常常被宗教、民族、艺术或个人等的偏见所主导。结果，人们所收获的只不过是一堆故事。20世纪初的时候，情况突然有了转机：爱因斯坦向我们展示了如何找到所有与物理定律和引力特性相一致的可能的宇宙、如何重构这些宇宙的历史，以及如何预测这些宇宙的命运。但实际上，要找到这些可能的宇宙谈何容易。从那时起，为了找到这些可能的宇宙，天文学家、数学家和物理学家都在拼命求解爱因斯坦那复杂难缠的方程组。这本书所要讲述的正是开拓者们披荆斩棘的故事，以及在他们面前徐徐展现出来的美妙画卷。

这是多么波澜壮阔的一幅宇宙图景啊！许多宇宙的名字取自20世纪最著名的一群科学家。有些宇宙膨胀，有些宇宙收缩。有些宇宙像陀螺一样转动，而有些宇宙则是一团乱麻。有些宇宙是绝对均匀的，而有些宇宙却是疙疙瘩瘩，或者在各个方向上被宇宙潮汐扯来扯去。有些宇宙始终在振荡，有些宇宙是一片冰冷的不毛之地，还有些宇宙失控了，膨胀速度越来越快。有些宇宙允许时间旅行，可以回到过去，而有一些宇宙则允许在一段有限的时间内发生无穷多的事情。仅有少部分宇宙能够演化出生命，其他宇宙仍是意识无法染指的处女地。有些宇宙会在一场爆炸里终结，有些宇宙会在哽咽啜泣里终结，而有些宇宙永远不会终结。

我们的故事会讲到物理定律可以随时间和地点不同而发生变化的宇宙、拥有额外隐藏的时空维度的宇宙、永恒的宇宙、位于黑洞之中的宇宙、毫无预兆就突然终结的宇宙、碰撞的宇宙、暴胀的宇宙，以及由其他东西变来的宇宙——甚至从虚无之中冒出来的宇宙。

渐渐地，我们会遇到关于今天我们所看到的宇宙最新最精确的描述，还有现代物理学理论引入的“多重宇宙”概念——包含所有可能宇宙的大宇宙。这些是人类一切的科学研究中最匪夷所思也是意义最为深远的思考。在我们的宇宙陈列馆中所展出的可能宇宙都真实存在吗？还是说其中只有一个宇宙获得了如此殊荣？

其他关于宇宙学和天文学的书籍把目光集中在了一些具体的问题

上，如暗物质、暗能量、宇宙的开端、暴胀、生命出现的巧合或是宇宙的终结，但本书将给读者呈现的是各种各样的宇宙及其发现史，以及发现了这些宇宙的科学家们的独特个性，所有这些内容会有机地结合在一起。

我要特别感谢 Katherine Ailes、Allen Attard、Donato Bini、Arthur Chernin、Hyong Choi、Pamela Contractor、Cecile De Witt、Charles Dyer、Ken Ford、Carl Freytag、Gary Gibbons、Owen Gingerich、Jörg Hensgen、Bob Jantzen、Andre Linde、Kay Peddle、Arno Penzias、Remo Ruffini、Doug Shaw、Will Sulkin、Kip Thorne、Don York等人在图书编辑、图片收集以及一些重要历史细节的澄清上提供的帮助。我还要感谢伊丽莎白的悉心照料，感谢两个已不再年轻的孩子提出的问题。还要感谢我的孙女，这本书就是献给她的。

约翰·D.巴罗

于剑桥

有些我们不知道的东西在做着我们不知道的事情。

——阿瑟·爱丁顿（1882～1944，英国天文学家）

目录

第1章 在正确的时间和地点

我知道这一切只是存在于我们的大脑中，但大脑是很厉害的。

——科林·科特里尔（英国作家）[1]

二人行

如果一个年轻人说他想研究宇宙学，我总是会觉得吃惊；我认为宇宙学是一种可遇而不可求的东西。

——威廉·H.麦克里（1904～1999，英国天文学家）[2]

老绅士沿街走来，跟往常一样——一身标志性的、有些凌乱的波西米亚风格。这个在美国大街上慢慢走着的欧洲人，神色黯淡，目光深邃，却不太在意自己要去哪儿。周围有很多购物的顾客和逆流而来的迟到学生，他在其中礼貌地穿行时，总是引起当地人的注目。似乎每个人都知道他是谁，他却避开人们的目光。今天，他有了一个新同伴，身材高挑，体格健壮，衣着略显寒酸也不太整洁，但不是老绅士那种样子。一路上他们都在讨论着，走着聊着，毫不在意身旁商店的橱窗。老绅士一边倾听着，一边若有所思，有时眉头微微一皱；他年轻的同伴则在情绪高昂地推销着自己的观点，滔滔不绝，时不时地用力比划着。两个人都操着一口不太地道的英语，而且口音也不一样，听得出融会了各地特色。过马路时他们停了下来，在路肩附近等了一会儿，让车辆行人先过去。交通灯一变，他们继续过马路，没说话，同时都注意着交通灯、声响和周围的一举一动。突然，有了新状况。高个子开口说了些什么，手激动了起来。这时交通又恢复了运行，老人却没动，定在原地，无视眼前车流和匆匆行人。同伴的那番话已经完全吸引住了他的思绪。车流从两旁轰隆隆穿过，两人便孤零零地待在路中间，像个人肉做的安全岛。老人陷入了沉思，而年轻人则不断重复着自己的观点。最后，他们又回到了身旁川流不息的现实中，都忘了刚才打算去哪儿，老人静静地领着年轻人回到人行道上（就是一分钟前他们待过的地方），然后朝着他们来的方向，继续走着聊着，完全沉浸在对这个新观点的讨论中。

两个人一直在讨论的内容是宇宙。[3]地点是在美国新泽西州的普林斯顿，时间则是第二次世界大战期间。年轻人叫乔治·伽莫夫（1904～1968），被朋友们唤作“吉吉”，是个流亡到美国的苏联人。年长的那个人就是阿尔伯特·爱因斯坦。在过去三十多年的努力中，爱因斯坦为我们展示了如何用简单的数学来理解整个宇宙的行为。伽莫夫则发现，宇宙必然拥有一个与现在大不相同的过去。让他们定在路中间的是伽莫夫的提议，即物理规律能够描述某些东西从虚无中创造出来。这可以是一颗恒星，但也可以是一个宇宙。

有趣的宇宙

历史就是那些本可避免的事情的总和。

——康拉德·阿登纳（1876～1967，联邦德国首任总理）

宇宙到底是什么？它从哪儿来？将向哪儿去？这些问题看似简单，却又是史上最深刻的一类问题。随着知识面的不同，我们口中的“宇宙”的含义也千变万化。[4]它不就是你能在空间中看到的一切吗——精确地说，也许再加上周围的空间？或者是一切有形的存在？当你列举出所有的事物，要把“一切”都囊括时，你会开始琢磨那些被物理学家称为“自然定律”的“东西”和其他一些诸如时间和空间之类的无形之物。尽管看不见摸不着，但你可以体会到它们的作用，它们看起来很重要，看起来是存在的（有点儿像足球比赛规则），所以我们最好也把它们列进去。那么要不要考虑过去和未来呢？仅仅局限于目前所存在的事物似乎有些狭隘。如果我们把过去曾存在的一切都当作宇宙的一部分，那为什么不把未来的也加上呢？最后我们得到的似乎是这样的定义，宇宙就是曾经存在、现在存在以及将要存在的一切。

若是我们再钻一下牛角尖儿，还可以把宇宙看得更恢弘一点，不仅包括可以存在的一切，还包括可能存在的、但还未发现的一切——最后，甚至还要包括不可能存在的一切。一些中世纪的哲学家[5]曾被这种完备性吸引住了，试图把这所有的一切列出一个过去、现在及未来可能存在或不可能存在的万物清单。本来已经问题成堆，这么做无异于雪上加霜。但是最近，现代宇宙学的研究中又出现了这种做法，尽管面貌有点不同。现代宇宙学不只研究我们宇宙的结构和历史，还关心别的可能存在的宇宙。我们的宇宙有大量特殊的、匪夷所思（至少对我们来说）的性质，这让我们不禁想估量一下，如果宇宙不是这样又会如何。这意味着我们不得不学会造出“别的”宇宙的例子，以便进行对比。

现代宇宙学就是这么回事。它可不仅仅只是试图尽可能完善、精确地描述我们的宇宙，它还试图将这种描述置于更广阔的种种可能性的语境当中。它在诘问，“为什么”我们的宇宙是这样而不是那样。当然，最终我们也许会发现，除了所看到的之外，再没有别的可能存在的宇宙（有着与我们设想中不同的结构、物质组成、物理规律和年龄，等等）。一直以来，宇宙学家们等待着甚至是盼望着这样的事情发生。然而最近的潮流却反其道而行之，使我们不得不面对种类繁多的可能的宇宙，并且它们都符合自然定律。更有甚者，这些别的宇宙也许并不仅仅限于一种可能性，它们也许能够像我们所谓的通常事物那样存在，就像此时此地的你和我一样。

地点很重要

他［雅各］做梦，看哪，一个梯子立在地上，梯子的顶端直伸到天；看哪，上帝的使者在梯子上，上去下来。

——《旧约·创世记》[6]

人们谈论宇宙的历史已有几千年。当然，那是他们认识的宇宙，可

别跟我们的混为一谈。对很多人来说，宇宙只不过是他们所能到达的最远的地方，或者是用肉眼就能看到的繁星点缀的夜空。大多数古代文化都试着为他们看到的世间万物找一个说法或编一个故事，无论是在天上、地下还是海里。[7]他们之所以描绘这种宏大图景，并不是因为对宇宙学有兴趣，而是为了说服自己或是他人相信事物都有其存在的意义，同时他们本身也是意义的一部分。如果承认真实世界的一部分是人类无法定义、无法掌控的，那么就会带来一种危险的不确定性。这也就是为什么在古代神话里，宇宙的本质看起来总是那么圆满（万物各有其位，各安其位）的原因。没有“大概”，没有注意事项，没有不一定，没有进一步探索的可能性。这些说法真的就是“万有理论”，只不过它们跟科学沾不上一点边。

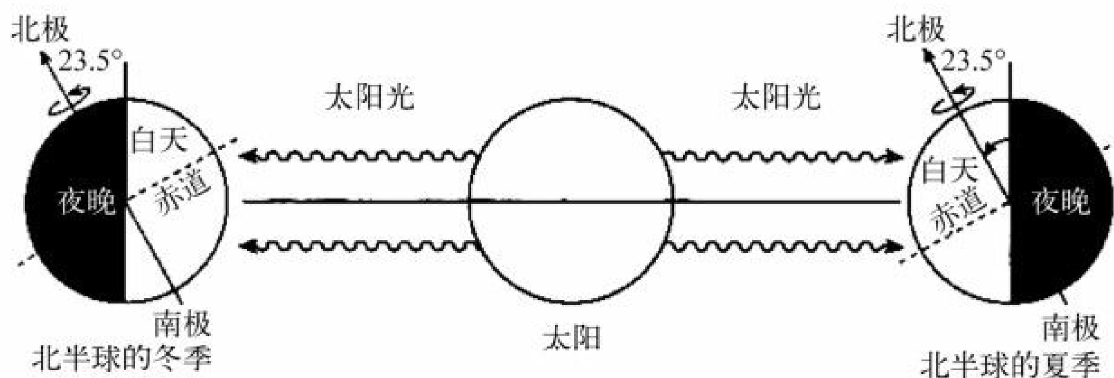


图1.1 地球的自转轴，一端是北极，一端是南极，并且与地球公转轨道平面的垂线相差23.5度

生于何时、身在地球何处，会影响你对宇宙的认识。如果你住在赤道附近，那么每天晚上恒星的运动就一目了然。它们初升，高高挂起，在你的头顶彻夜闪耀，徐徐下落直到没入另一侧地平线。夜复一夜，你就像是天体运动的中心。可如果你离热带很远，天空看起来就大不一样了。夜里有的星星东升西落，越过你的头顶，而后又回到地平线的怀抱。另外一些星星则从不落下，总是挂在地平线之上。它们似乎在围绕空中的一个大圆心转圈，就像被钉在了一个绕轴转动的轮子上。这必然会引起你的好奇心，群星围绕的那一点究竟有什么特别的呢？为了解释这些夜夜旋转的群星，北半球的居民们创造了许多关于那个天空大路标的传说和神话。

世界各地的星空之所以呈现出如此的差异，是因为地球的自转轴是倾斜的（图 1.1）。当地球绕太阳公转时，贯穿南北两极[8]的地球自转轴

并不垂直于其公转轨道，而是倾斜了大约23.5度。这造成了不少显著的影响，四季变化就是这么来的。假如没有这个倾角，就不会有每年的四季变化；假如倾角过大，季节的更替就会剧烈得多。不过，即使你完全不知道地球在绕太阳公转，也不知道自转倾角，但只要每晚观察一下星空，这个倾角让不同纬度下的天空呈现出了不同的模样。

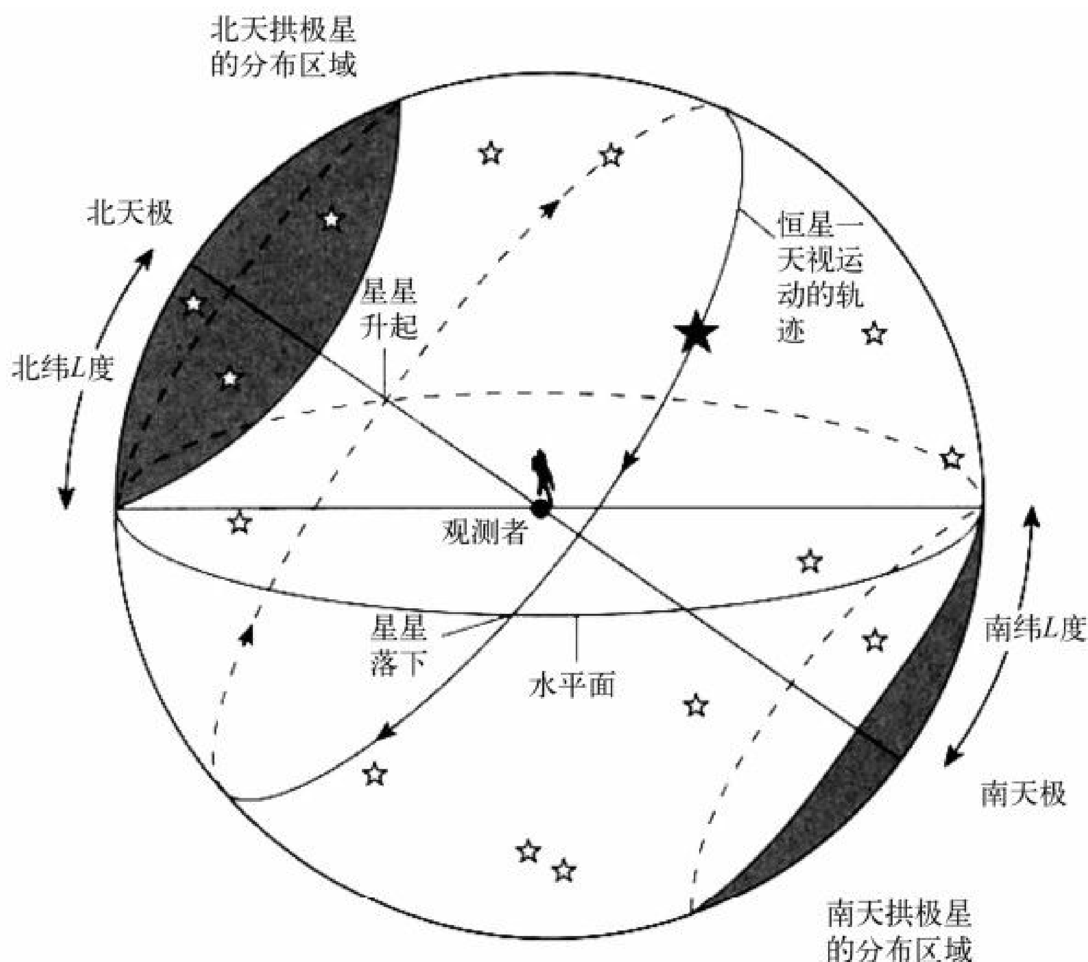


图 1.2 生活在北纬 L度的天文学家所看到的星空。无论何时，他们都只能看到半个星空。有些星星，例如靠近北天极的拱极星永远不会落入地平线。而另外一些星星，例如南天极周围的拱极星，永远都不会从地平线上升起，所以北纬L度的天文学家就没法看见

如果我们把南北两极的连线延长到太空中，所指示的方向我们就称之为南天极和北天极。当地球自西向东自转时，夜晚静止的恒星看起来就像在沿着相反的方向扫过天空。如果白天也能看见它们的话，地球每完成一次自转，它们也就在天空中画完了一个大圆圈。尽管如此，天上

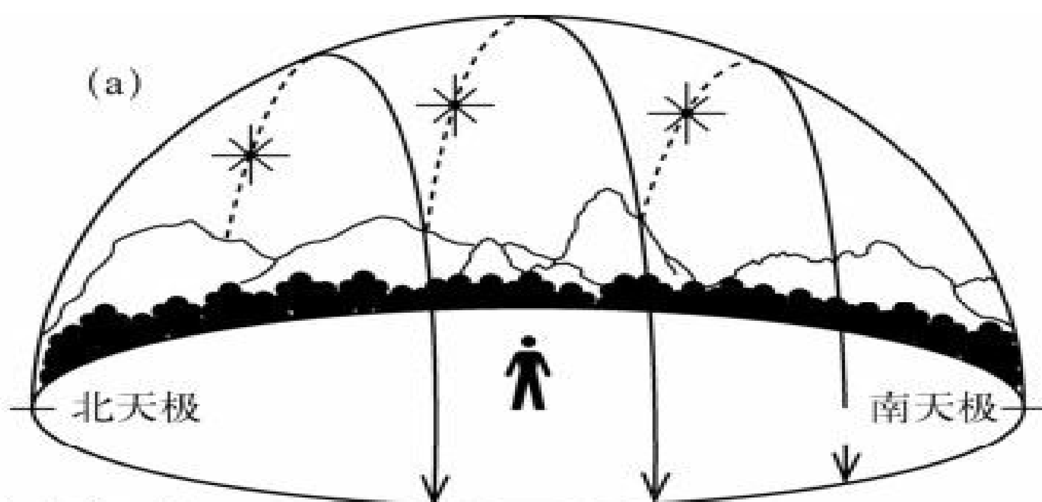
的圆圈并不是全部都看得见，因为圆的一部分可能在地平线以下。图1.2显示了北半球北纬L度的人们在晴朗的夜晚所能观察的景象。[9]

地平线把每个观察者的天空分成两半，任何时候都只能看到地平线之上的天空。在北纬L度的地方会看到北天极位于地平线之上L度，而南天极在地平线之下 L度。地球自转使得天空像是在自东向西地围绕着北天极转动。于是我们看到星星从东方的地平线出现，渐渐升起，到达各自的最高位置，或者叫“天顶”，然后落向西方的地平线。[10]

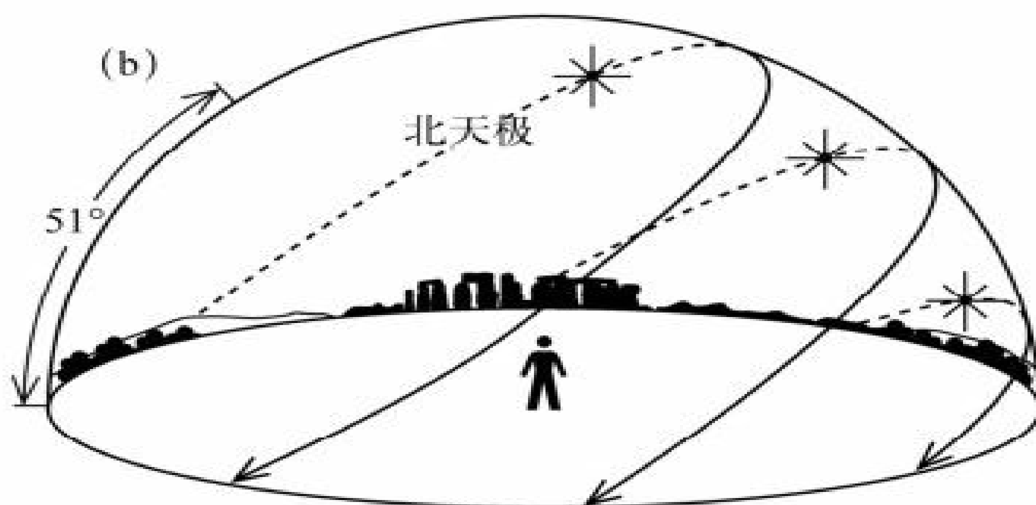
有两类恒星没有这样每夜上升下落。在北天极附近L度的圆盘范围内，恒星能在天空中绕一个完整的圆，从不会消失在地平线以下。如果夜空足够暗又没有云，它们就一直看得见。[11]对于目前的欧洲天文观测者来说，这个范围包括大熊座和仙后座的恒星。另一方面，对于图中北半球的天文观测者来说，在南天极附近相同范围内的恒星是永远也看不到的。在他们眼中，这些恒星从来没有从地平线上升起过。[12]这就是为什么地处北半球的欧洲永远也看不到南十字星座。①至关重要的是，随着观察者所处纬度不同，这些永远可见和永不可见的星空区域的大小也会变化。当你远离热带，纬度渐渐变高时，这些区域的范围也在扩大。图1.3显示了地球上三个不同纬度地区的观察者所看到的星空。

① 南十字星座是全天88个星座中最小的一个，每年4、5月间我国北纬20度以南地区的地平线附近可以短暂地看到。澳大利亚、新西兰、巴西等国的国旗上也有南十字星座。——译者注

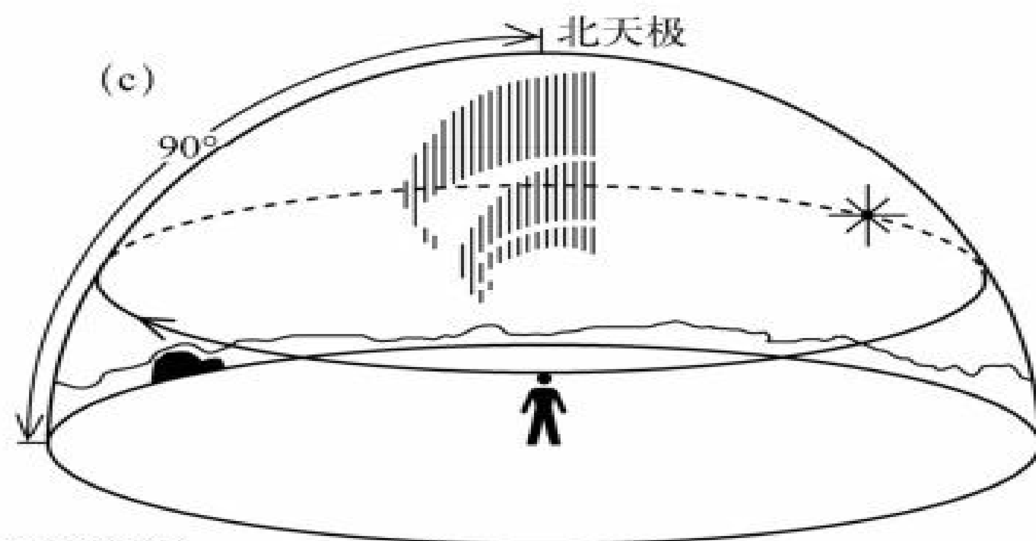
赤道的纬度是 0度，所以夜晚的星空之中不存在永远可见或永不可见的区域。赤道上的天文观测者可以瞥见天上每一颗亮星，不过实际上，两个天极会被地平线附近的尘霾所掩盖。群星升起，上升到各自在天空的最高点。每一颗恒星升起后，运动方向都保持不变，于是可以成为整个夜里户外旅行者和航海者们的绝佳导航灯塔。几乎没有一点儿侧向偏差，天空看起来非常简洁、非常对称。观察者们会觉得自己处在一个拱形天幕中心的正下方，一目了然的运动像是专为他们而设计的。他们犹如宇宙的中心。



赤道（纬度为0度）



巨石阵（北纬51度）



北极（北纬90度）

图1.3 人们所生活的纬度不同，看到的星空也不同：(a)在赤道看到的星空，(b)在英国巨石阵看到的星空，(c)在北极看到的星空。产生这种差异的原因是，群星围绕的天极所处的位置不同



图1.4 这张照片经过了对北天极方向的长时间曝光，北天极正好在树顶上，处于图片的中心，周围是群星的运动轨迹

北极是另一种极端情况。那里的纬度是90度，天上的恒星既不会升起也不会落下。它们会绕着天空不断转圈。北天极就在头顶，被群星环绕着。那就像是宇宙的焦点，而我们正好位于它的下方。

在一些没那么偏北的地方，例如北纬51度的不列颠远古巨石阵，情况就介于两个极端之间。北天极附近51度范围内的恒星的运动轨迹围着北天极形成了完整的同心圆。其他的恒星则从地平线升起，到达天顶，而后落下。天空呈严重地一边倒的态势。恒星们升起落下，各走各的轨道。最醒目的是，盘旋在极点周围的恒星都围着它转，好似宇宙大转盘的轴心（图 1.4）。对那些不懂天文学，也不知道地球在转动的观察者们来说，天上似乎存在一个特殊的位置。

这就是为什么关于天空的神话和宇宙的本质，不同地区有不同说法的原因。在远离赤道的斯堪的纳维亚和西伯利亚，我们能找到关于天空大圆圈的传说：众神就住在大磨盘的中心。最靠近天空大圆圈中心的那颗星有着非同寻常的地位，它领导着宇宙大帝宝座周围的群星阵列。[13]

在这里我们不打算追溯这些神话。我们只是想强调，当人们受到地域局限的时候，想要提出一种解释宇宙的图景是多么艰难。如果不了解星空，不清楚地球的转动和方向，你会不知不觉地蒙上强烈的偏见。

即使最灿烂的早期文明，当它开始观察天空时，也不免受到了地域局限性的影响。我们被限制在一颗行星上，同别的行星一起绕恒星转动。如今我们知道除了太阳系有行星之外，有数百颗遥远的恒星也有围绕自己的行星（目前是五百多颗①）。熟悉了这些之后，我们就容易跳出地球中心论的框框，从而理解世界上还有其他的行星。从一个简单的例子可以看到跳出这个框框的难度，那就是在地球上观察其他行星的运动，例如火星。我们假设地球和火星的公转轨道都是圆形，而且火星的轨道半径大约是地球的1.5倍（图1.5a）。地球公转一周需要1年，假设火星的公转周期是 2年。现在计算出两者轨道之差随时间的变化。这就能告诉我们，从地球看去火星的视运动是什么样子。图1.5b显示的就是这样一个示意图。

① 截至2012年底，科学家已确认发现的太阳系外的行星总数达八百多颗。——译者注

这个古怪的心形曲线（数学上叫蚶线）有一个环，很有意思。从曲线的最右端向左移动时，我们看到火星离开地球而去。当曲线下降到与水平坐标轴相交的-5处时，意味着两颗行星分别在太阳的两边，相距最远。当火星开始向地球靠近时，奇怪的事就发生了。火星离地球越来越近，貌似就要撞上去了。它却突然调转方向，拉开距离，重新进入远离地球的漫长周期。在火星离地球很近的那段时间里，如果你连续数个晚上用肉眼观察星空，就可以发现这个“逆行”现象②。我们会发现，每当两颗行星靠近最近点时都会发生这种现象。如果我们换一个更远的对象看，例如公转一周需要29.5年的土星，在每个土星公转周期内，地球和土星的相对运动靠近最近点的情况会发生好几次，所以在对应视运动的图上会有好几个环。[14]

② 火星逆行的视运动轨迹图片，可参见：http://www.phys.ncku.edu.tw/~astrolab/mirrors/ apod/ap080511.html. ——译者注

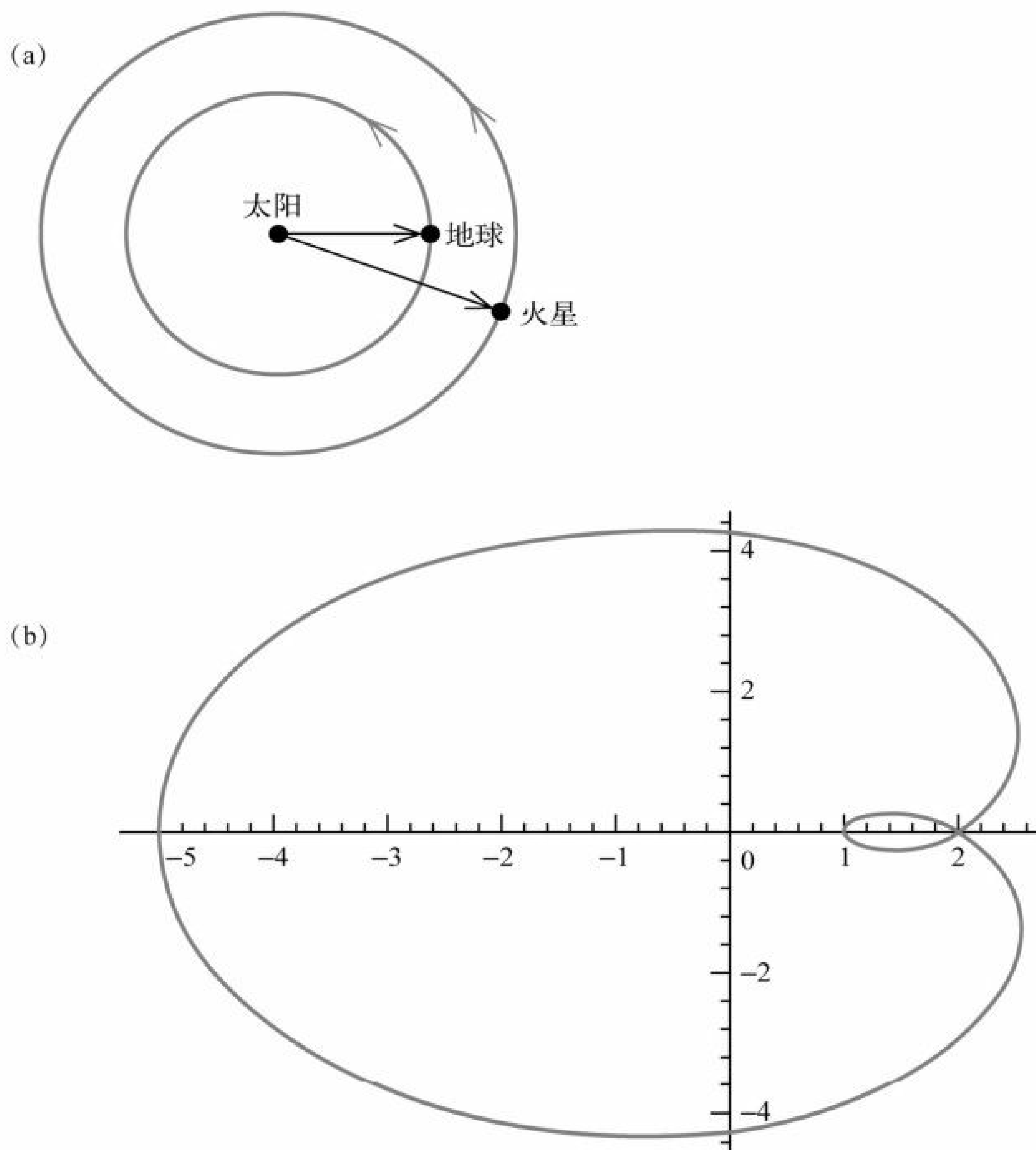


图1.5 从地球上看到的火星视运动轨迹。(a)假设地球和火星的轨道都是圆形的，火星的轨道半径大约是地球的1.5倍，火星公转一周大概需要2年左右（687天）。(b) 从地球上看，火星的运动形成了这种闭合的心形轨迹，叫做蚶线。起初，火星离地球越来越远，最远时相距2.5倍的地球轨道半径，此时地球和火星分别位于太阳的两侧。此后火星开始朝着地球靠近，到达最近点之后突然调转方向往回走。然后它再次调转在天空中的运动方向，开始渐行渐远

对此我们得到的教训是，如果没有大局观，没有一个运动学理论，就很难解释那些天上的运动。如果早期的天文观测者观察火星两年，会看到火星离我们而去，又返回来，接着被明显地推了一下，再次离我们而去。是什么样的力在起作用？如果我们待在地球上，又不知道所有的行星（包括我们所在之处）都在以不同的速率绕太阳旋转的话，那么这个问题就很难回答。

亚里士多德的球对称宇宙

所谓专家，就是那种小心避免了小错却大步迈向大错的人。

——本杰明·施托尔贝格（Benjamin Stolberg，1891～1951，美国作家、记者）

约公元前350年，亚里士多德（公元前384年～公元前322年）提出了一种哲学上的宇宙观，试图对这些现象加以简化，由此一幅解释天体视运动的复杂图景逐渐浮现。亚里士多德猜想，世界并不是从过去的某个时刻突然开始的，而是在过去和未来都一直存在，本质永远保持不变。他把对称性看得非常重要，并相信球体是世间最完美的形状。因此，宇宙必然是球对称的。为了描述天空中看得见的天体和它们的运动，亚里士多德提出一种复杂的洋葱皮结构，包含不下55层透明水晶嵌套的球面，这些球面都以地球为中心。亚里士多德又假设，地球也是一种球体（这个假设很难与当时他所能看到的情况相符）。每一个看得见的天体都固定在其中一个水晶球面上，以不同的速度绕中心旋转。除此之外，还有许多别的球面带着行星一起运动。这样，亚里士多德就既能解释观测到的现象，又可以预言还未被看到的新事物。这套理论拥有一个现代科学理论的众多特点，当然也有众多大不相同之处。在亚里士多德的图景中，恒星所处的外层球面是物质无法存在的区域——一个精神领域。我们看到的所有运动都始于第一推动者，在这个域的边界上发力，使外层球面转动。于是，转动会一层一层向内传递，直到整个宇宙处于完美的转动之中。对不同球面的转动速度进行修修补补之后，就可以解释夜空的许多特征。

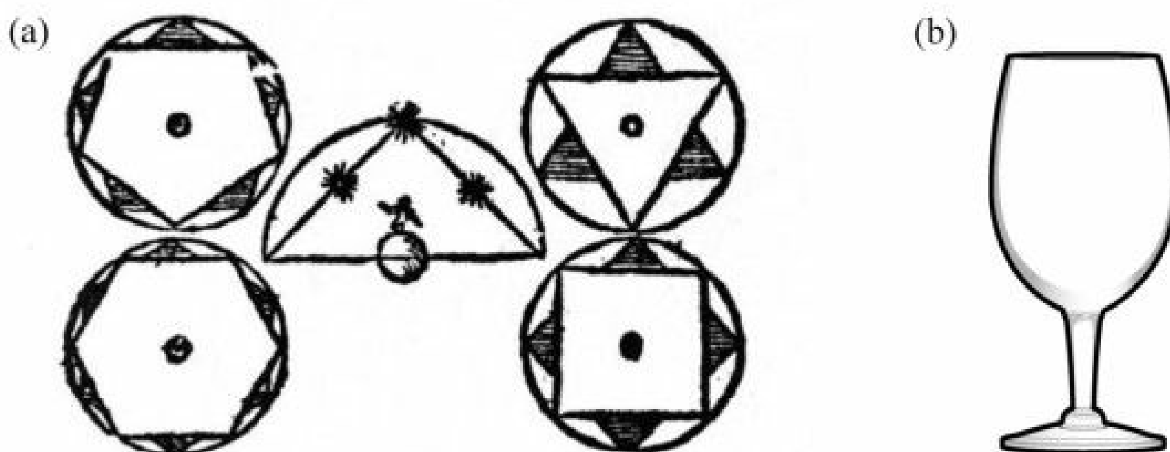


图 1.6 (a) 一个旋转的球面总是在空间中占据相同的地方，但多面体旋转时就会制造出“真空”。于是，这种亚里士多德式的“证明”便得出了大地是球形的结论，图片来自罗伯特·雷柯德（Robert Recorde， 1510～1558）的《知识的城堡》（Castle of Knowledge，1556）一书。不过，一个轴对称的酒杯形状的宇宙，如(b) 所示，仍然满足亚里士多德的条件，也就是转动的时候不会产生真空

后来，中世纪的基督教学者吸收和改造了亚里士多德的观念。他们认为，第一推动者就是《圣经·旧约》中的上帝，而最外层的球面就是基督的天堂。地球中心说与中世纪时以人类为中心的世界观正好相呼应。

地球和其他所有球面的球对称形状都有一个重要的特点，那就是球面转动时不会切入没有物质的虚空中，也不会在身后产生虚空（图 1.6a，它由英国都铎时期的杰出数学家、医生罗伯特·雷柯德所画）。真空不可能存在。如果说无穷大的物理量不可能存在，那么真空就更不可能存在。[15]地球永远保持球对称的形状，并且在转动时，总是占据相同大小的空间。如果地球是立方体就不行了。[16]实际上，在亚里士多德的论述中，球体并不是唯一能使地球不进入或不产生虚空的形状。葡萄酒杯的形状也可以。[17]

亚里士多德并不像我们（遵循牛顿的观点）理解引力时那样，把运动看成是不同天体之间作用力的结果，而是把力看成物体的固有属性。它们以一种“自然的”方式运动。圆周运动是最完美最自然的一种运动。

托勒密的**“希思·罗宾逊”**①宇宙

① 威廉·希斯·罗宾逊（William Heath Robinson，1872～1944）是一

位英国漫画家，以绘制古怪的机器闻名。在英国，“希斯·罗宾逊”一词被用来指代设计得复杂异常、实现的功能却非常简单的发明。——译者注

> 我曾是个天文学家，只不过我总是轮到白班。
>
> ——布莱恩·马洛（自称是科学喜剧演员）[18]

我们在前面说过，太阳是太阳系的中心，不同的行星以不同的速率围绕太阳转动。在这种情况下，你就会观察到天空中存在一种奇怪的运动，会看到别的行星在一段时间逆着平时的方向运动。这其实是一种幻觉，是因为我们同其他行星之间存在相对运动。不同的行星都在以不同的角速度转动，所以有时我们会看到其他行星表现出反常的逆向运动。亚里士多德和他的追随者需要解释这些现象。

克罗狄乌斯·托勒密（约 90～168）在公元 130年左右首先发现了这个挑战性问题的一个解决方法。这个方法是古代最接近“万有理论”的东西，并且延续了一千多年的时间。托勒密面临的挑战在于，他需要把行星的复杂运动，包括所有的逆行运动，同亚里士多德认为的地球是宇宙的中心的严格观念结合起来。其他所有天体都要在圆形轨道上以不同的角速度绕地球转动，而且宇宙中的任何天体都不能改变亮度或是其他的内在属性（图1.7）。这件事确实很有挑战性。

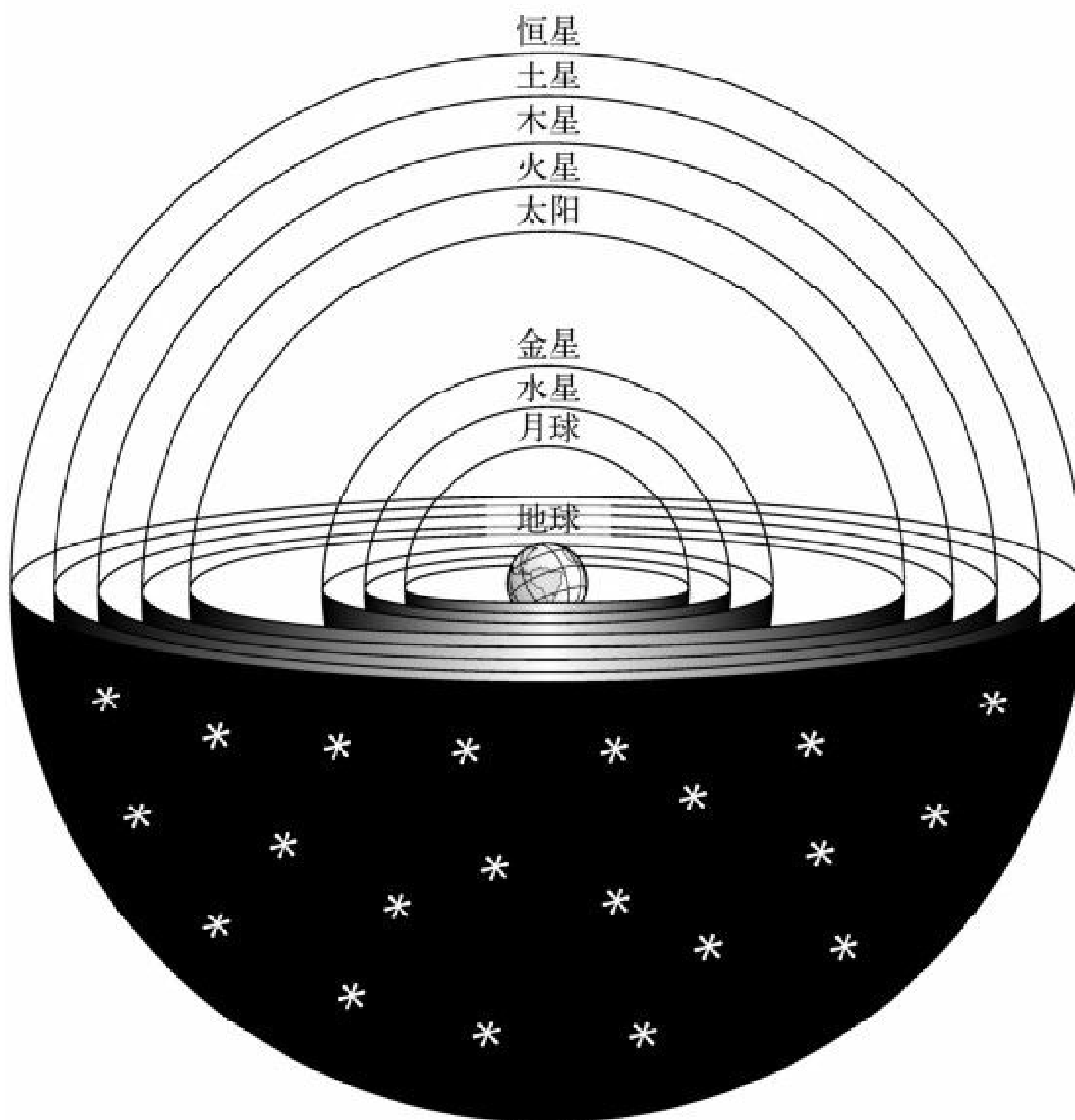


图1.7 亚里士多德和托勒密的宇宙模型

托勒密在他的著作《至大论》中是这样回答这个难题的：从行星或太阳绕地球的圆形轨道（或叫做“均轮”）可以反推出一个点的运动，而这个点又是该行星的另一个小型圆周运动（又叫做“本轮”）的圆心，行星就沿着本轮运动。[19]行星的总体运动看起来就像一个由绵延不断的螺旋线套着的圆（图1.8）。

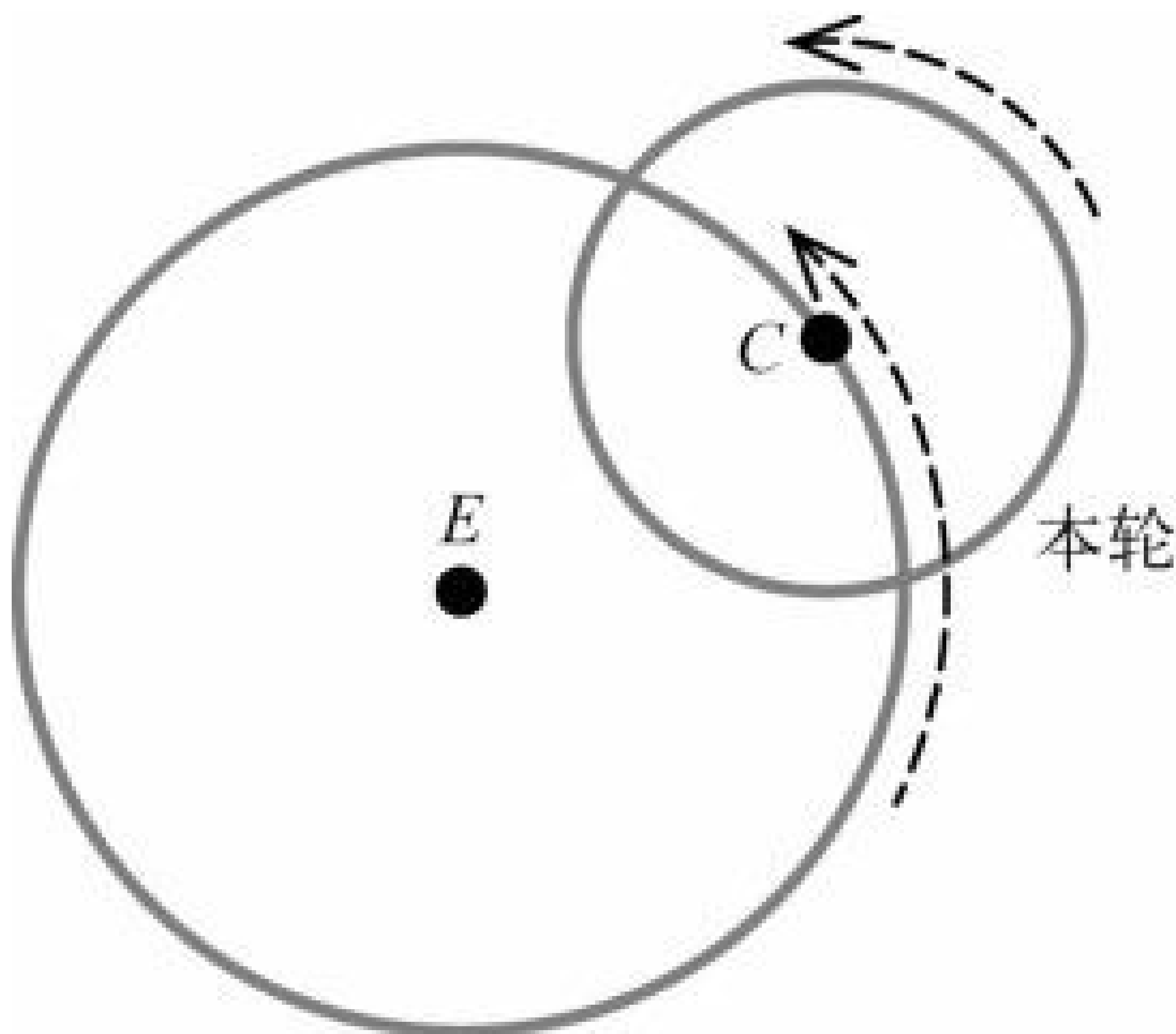


图 1.8 本轮。一个行星沿着一个小圆运动，也就是本轮，而本轮的圆心C，再沿着一个大圆运动

从地球上看，像火星那样的行星的总体运动是绕着某个动点的圆，而那个动点本身又处在一个围绕太阳的圆形轨道上。随后托勒密精益求精，又在行星绕太阳的轨道上增加了更多本轮（做圆周运动的圆）。为了进一步追求更高的精确度，中世纪的哲学家们继承了托勒密的方法，增加了越来越多的本轮。[20]

以非常高的精确度把行星和太阳的所有运动特点与观测对应起来，又可能会改变许多别的东西。增加本轮可以很好地描述地球上看到的逆行运动。在行星绕小本轮的轨道上，有一半时间行星的运动方向与它在均轮的运动方向相同；而另一半时间则是相反的，我们就看到了逆行。从地球上看，那个行星就会突然减速，在天空中停住，然后反向，再次

减速，再次停住，再次反向。这是真正的逆行，而不是由行星以不同速率绕太阳转动而产生的假象。

关于行星和太阳绕地球的复杂运动，古代人所给出的回答表明了仅从观测入手，或从一般的哲学原理出发，要想正确地描述宇宙是多么艰难。如果当初亚里士多德有进一步追问，那他就不得不努力解决另一些尴尬的问题。为什么地球并不是完美的球形？既然地球的中心地位被看得这么重要，那为什么其他圆周运动可以在本轮上进行，而本轮的圆心并非地球？为什么人们能接受每个行星均轮的圆心都偏离了地球的想法？偏离得可能并不多，但要么地球是宇宙的中心，要么就不是啊。

哥白尼革命

如果全能的主在创造万物之前咨询我的意见，我肯定会建议让某些事物变得简单些。

——阿方索十世（1221～1284，卡斯蒂利亚王国国王）[21]

托勒密的地球处于宇宙中心的模型，是人类所创造的一种复杂构想。它并不正确，可是有太多办法能够对它进行微调以符合对于行星别扭运动的最新观测结果，因此它一直延续了下来，并没有受到太大的挑战。这种情况一直持续到15世纪。托勒密理论灵活多变的特点导致“本轮”一词成了专门用来描述不靠谱的或过于复杂的科学理论的贬义词。如果你构建一个理论的时候不断往里面添加细节，试图解释所有可能发生的新现象，那你的理论就不会有什么说服力。就像你提出了一个汽车理论并预言说，所有的汽车都是红色的。星期一早晨你出门后发现了一辆黑色 汽车，于是你就修改了理论，然后预言说除了星期一那天有一些黑色汽车外，其他所有的车都是红色的。许多黑色和红色的车在面前开过，一切看起来都不错。但在下午又开过来一辆绿色的车。好吧，除了午后有绿色的汽车之外，所有星期一的汽车都是红色或黑色的。你能看出来接下去会如何。这就是一种有一系列“本轮”修正的汽车理论。每一个新现象都能被微小的修正所解释，这样就可以保住你最开始提出的大前提。但到了一定时候，你就该意识到必须要推倒重来了。

这个例子当然有些夸张，托勒密的理论则要老练得多。每加入一个新的本轮，就相当于引入一个新的小修正，用以解释人们观测到的行星运动的精确细节。这个理论是第一个将收敛近似过程应用于实际的例子。每次对模型的修正都比上一次的微小，并且比从前的理论更符合天文观测的结果。[22]对大多数情况来说，这种方法相当好用。但是整个太阳系的图景全错了，太阳系的中心也弄错了。而对此加以辩驳的是一个非常有说服力的理由。

尼古拉斯·哥白尼（1473～1543）一般被看成是一位革命者——那个将人类从宇宙中心位置废黜的科学家。但其实真相更为复杂，也更缺少戏剧性。如果他还算个革命者的话，那也只能算是个不情愿的革命者。[23]1543年，哥白尼去世后不久，他的著作《天体运行论》就寄给

了出版商。这本书没有掀起什么波澜，印数不太多，读的人就更少。然而，哥白尼的远见卓识还是及时地将人类的宇宙观引向了一个转折点。最终，它取代了托勒密的行星理论，把地球中心说换成了现在众所周知的太阳中心说。[24]

16世纪早期的印刷技术有所提高，这就意味着在哥白尼的书中可以把图表嵌入到相关文字中去。其中最著名的一幅图（图1.9）描绘的是一个太阳处于中心的简单模型。最外面的一圈表示我们太阳系之外，也就是“恒星的静止球面”的边界。其他六个圈表示当时已知的六颗行星的运动轨迹。从外向内看，分别表示的是土星、木星、火星、地球（旁边的新月表示月球）、金星和水星。所有行星都在圆形的轨道上绕太阳（Sol）转动。月球被认为在圆形的轨道上绕着地球转动。

哥白尼和托勒密的体系并不是16、17世纪时仅有的关于太阳和行星的图景。图 1.10，摘自乔万尼·里奇奥利（Giovanni Riccioli，1598～1671）于1651年所著的《新至大论》[25]，很好地总结了摆在后哥白尼时代的天文学家面前的各种世界图景。这里有六个不同的太阳系模型（标为Ⅰ~Ⅵ）。

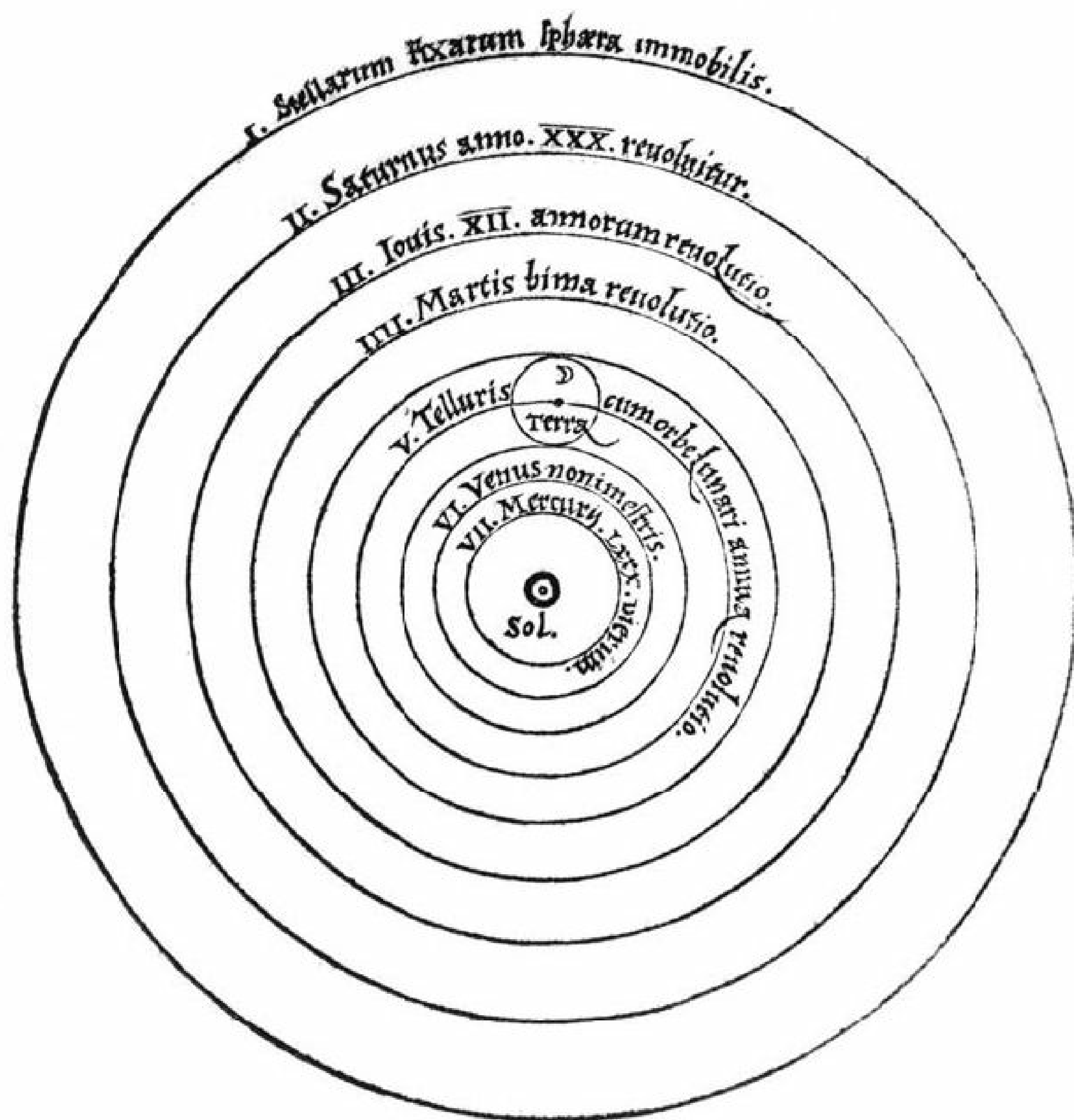


图 1.9 哥白尼的太阳中心模型，发表于 1543年。太阳周围的同心圆用罗马数字进行了标记。最外面的恒星层（Ⅰ）是固定的，里面是转动的行星层（Ⅱ~Ⅶ），依次是土星、木星、火星、地球（月球被表示成了新月的形状）、金星和水星的轨道

模型Ⅰ是托勒密体系，地球处于中心，太阳的轨道在水星和金星轨道之外。

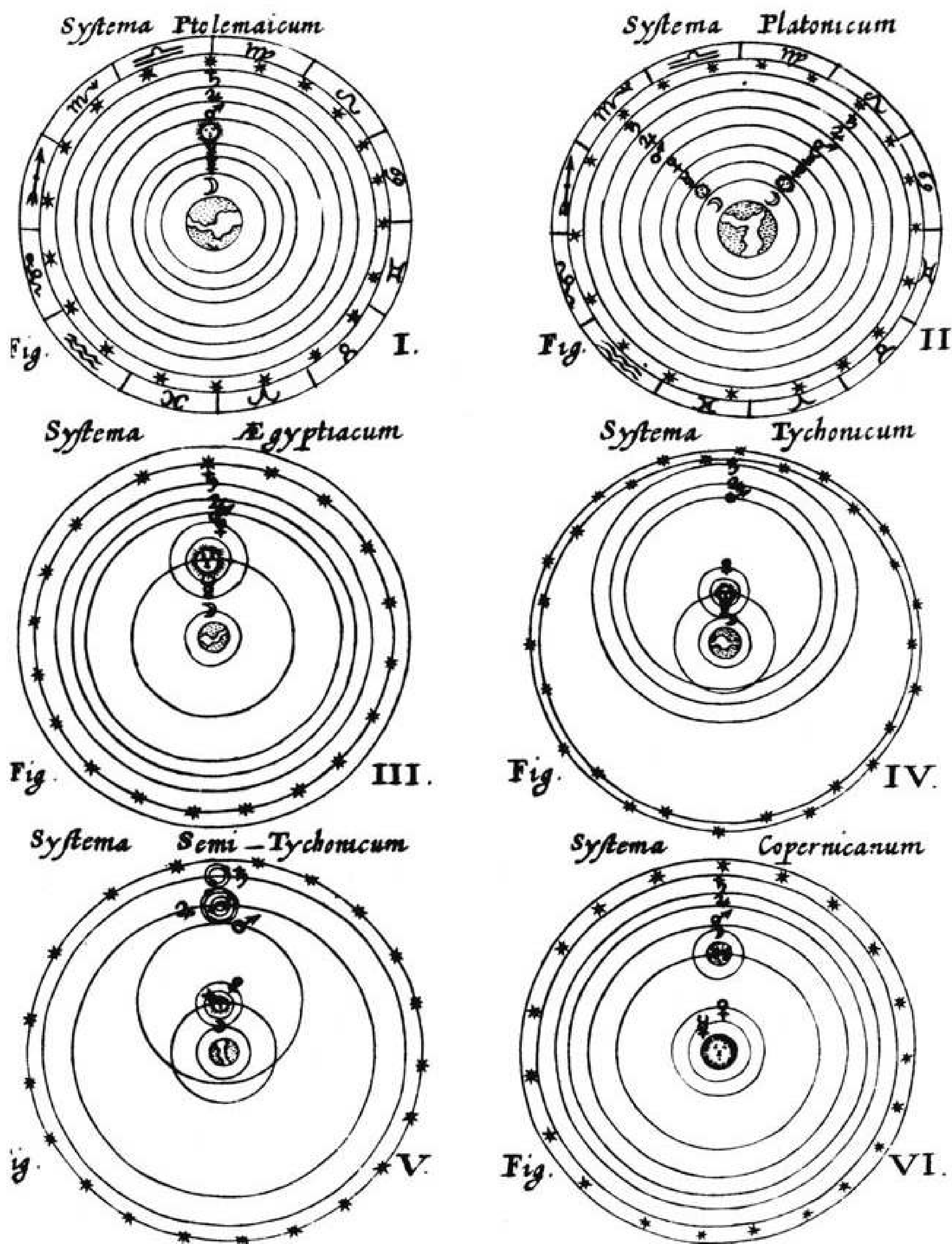


图1.10 乔万尼·里奇奥利在1651年出版的《新至大论》中展示的六种主要的世界体系

模型Ⅱ是柏拉图体系，地球在中心，太阳和其他所有行星都在环绕地球的轨道上，但太阳的轨道在水星和金星轨道的里面。

模型Ⅲ称为埃及体系，其中水星和金星绕着太阳转，并随着外面的行星一起绕着地球转。

模型Ⅳ是伟大的丹麦天文学家第谷·布拉赫（1546～1601）提出的第谷体系，其中地球被固定在中心，月球和太阳绕着地球转动，但其他所有行星都绕着太阳转。因此水星和金星的轨道有一部分在地球和太阳之间，而火星、木星和土星的轨道则整个把地球和太阳包含在内。

模型Ⅴ叫做半第谷体系，由乔万尼·里奇奥利本人发明。在这个模型中，火星、金星和水星绕着太阳，并随着木星和土星绕地球转。里奇奥利之所以想要把木星和土星同水星、金星和火星区别开，是因为当时已经知道它们像地球一样有卫星（火星的两个卫星当时还未被发现），所以它们的轨道必须以地球为中心，而不是太阳。

模型Ⅵ就是哥白尼体系，我们已经在图1.9中展示过。

众多的古代宇宙观教给了我们一些简单的道理。仅靠观察宇宙就想理解它并不容易。我们被局限在一个特殊行星的表面上，同其他行星一起环绕着一颗中年恒星。因此，我们在地球表面所处的地点和时间以及可能抱有的对于我们应在大千世界中处于什么地位的观念，都强有力地决定了我们从夜空中能看到什么。我们的宇宙观预先确定了我们的宇宙模型。

随着我们对于宇宙的视界日益宽阔，这些问题也变得更为重要了。为了有所进步，我们需要描述和预言我们所能看到的宇宙中的天体运动。最终我们想了解整个宇宙究竟是什么样子的。在这个方向上，18世纪的天文学家果断地迈出了第一步。让我们跟上他们的步伐。

注释

[1] C. Cotterill,*The Coroner’s Lunch*, Quercus, London, (2007), p. 123.

[2] Presidential Address, Royal Astronomical Society, February 1963, see W. H. McCrea,*Quart. J. Roy. Astron. Soc.* 4, 185 (1963).

[3] G. Gamow,My World Line, Viking, New York (1970), p. 150.

[4] 关于 universe一词的词源，可以追溯到 12世纪古法语中的univers，而这个词又来自从前拉丁语中的universum。而这个词的源头又是unus，意思是“一”，以及vesus，动词vertere的过去分词，意思是“转变，围绕，合为或改变”。所以，我们得到的字面意思是，一切都“变成一”，或者“合为一体”。在卢克莱修（Lucretius，约公元前 99年～约公元前 55年）创作于约公元前 50年的拉丁文诗歌《物性论》（*De rerum natura*）的第4章第262行中，人们首次发现了universum的诗歌缩写体unvorsum的用法。另一种关于词源的解释则主要与“一切都围绕一”的想

法有关，这个想法反映了古希腊宇宙学的思想，即天空最外层水晶球面转动，因而把变化和运动传递给了内层的行星球面，地球则静止地待在这些球面的共同中心。

[5] 这其中包括著名的爱尔兰哲学家约翰·斯考特·爱留根纳（Johannes Scotus Eriugena，815～877）。他将大自然（事物最全面的集合）划分为存在的和不存在的两组。这两组又可以进一步分为四类：（1）可以进行创造，但不能被创造的；（2）可以进行创造，也可以被创造的；（3）可以被创造，但不能进行创造的；（4）既不能被创造，也不能进行创造的。爱留根纳把上帝归为第一类和第四类，所谓一切事物的开端和终结。柏拉图式的理念世界属于第二类，而第三类则是上帝主宰的物质世界。在13世纪后期，中世纪的教会区分了三种“万千世界”或万千宇宙，一种是在时间上前后相继的万千世界，一种是在时间上同时存在的万千世界，还有一种是在空间上并存且被虚空隔开的万千世界。

[6] Genesis 28: 12–13.

[7] 世界的起源是什么？是否会终结？如果会，又将以怎样的方式终结？人类对这类问题的探索有着悠久的历史。这种探索通常最早是以神话故事的形式出现，也就是把宇宙看作一种“东西”。就像地球上环绕在我们周围的那些东西一样，宇宙也是由别的东西做成的。受到各自不同生活方式的影响，人们提出了不同的宇宙模型。有的模型认为，宇宙是一种活的东西，是由众神孵化出的。有的模型认为，宇宙是一个人潜入海底时捉到的。有的模型认为宇宙起源于自泰坦之间的一次冲突。还有一些模型认为，宇宙就像一粒种子，发芽生根，到时候它会枯萎，然后又重生，就这样永远处于生死轮回之中。有的文化认为，宇宙根本没必要存在一个开始的时刻。他们认为宇宙的生死轮回是永恒的，不存在什么东西都没有的情况：宇宙就是一种东西，所以不可能什么东西都没有。即使在后来基督教“从无中创造”的创世理论中，也不存在什么都没有的情况。上帝永远存在，即使物质的宇宙不存在的时候也存在。在古希腊的思想家，如柏拉图（约公元前427年～公元前347年）看来，在表面的现象背后，存在一些永恒的法则或理念。在古代，没有任何理论认为，宇宙会无缘无故地产生。但这是一种非同寻常的想法。我们通常认为周围一切事物的产生都有各自的原因。桌子有其产生历史，由于某种原因，它从不那么有序的木头变成了桌子。可是，宇宙是不是一种像桌子一样的“东西”呢？又或者，宇宙就像社会一样，是某种东西的集合？两者之间的区别非常重要，因为社会的每个成员都有自己的母亲（各自的原因），但社会本身并没有母亲。关于各种文化的创世神话，以下书籍的内容非常全面：M. Eliade,*The Myth of the Eternal Return*, Pantheon,

New York (1954); M. Leach,*The Beginning: Creation Myths around the World*, Funk and Wagnalls, New York (1956); C. H. Long, *Alpha: The Myths of Creation*, George Braziller, New York (1963); E. O. James,*Creation and Cosmology*, E. J. Brill, Leiden (1969); C. Blacker and M. Loewe (eds), *Ancient Cosmologies*, Allen and Unwin, London (1975).

[8] 地磁南极和地磁北极的连线并不平行于地球的自转轴。

[9] 伦敦在北纬51.5度，而新加坡在北纬1度。

[10]古代人把太阳在一年中所走出的圆形轨迹分成了黄道十二宫，并把它们叫做十二星座。古人认为地球是天体运动的中心，所以太阳绕着地球的视运动就会经过十二星座。人们在天空中给每个宫分配了大约30度长（十二宫加起来就是 360度的圆）、18度宽（约定俗成）的地方。

[11]这些星星叫做北天拱极星。

[12]由于天空中存在这些不可见的区域，而且随着观测者所处纬度和时期（就像陀螺一样，地球的自转轴正以 26 000年为周期在发生摆动）的不同，不可见区域的分布也会发生变化。人们试图利用这一规律来确定古代星图制作者所生活的时期和纬度。古代星图所描绘的星空和今天观察到的星空有所不同，这是因为对当时的观测者来说，今天我们可见星空的一部分区域永远都不会从当时的地平线上升起。这些研究还面临着诸多困难，最近的一篇综述请见：J. D. Barrow, *Cosmic Imagery*, Bodley Head, London (2008), pp.11–19.

[13]关于北方神话中大磨盘问题的调研，有一本书很有趣：Hertha von Dechend and Giorgio de *Santillana*, Hamlet’s Mill, Gambit, Boston (1969)。作者说，由于星空呈现为一个大磨盘的样子，北半球高纬度地区流传的天空神话之间存在某种文化上的一致性。不过，由于对某些问题一概而论，这本书招致了古代天文学研究者的批评，可参见以下书评：C. Payne-Gaposchkin, *J. Hist. Astronomy* 3, 206 (1972).

[14] D. G. Saari,*Collisions, Rings, and Other Newtonian N-Body Problems*, American Mathematical Society, Providence, RI (2005).

[15]实际上，在亚里士多德看来，这两个问题是一回事。他认为，绝对的真空不会产生阻力，因此物体的运动速度就会变成无穷大。大自然在任何情况下都不允许绝对真空和真正的无穷大出现，这种观念在西方哲学中延续了一千五百多年。

[16]这是人们第一次在物理学里提到拓扑问题。实际上，亚里士多德的论述并没有将球体理解为地球唯一可能的形状。为了不让地球转动时产生真空，或进入真空的区域，地球的形状只要是由任何一种旋转对称的图形绕着中心自转而成的就行了。

[17]任何由同心圆盘堆成的形状都可以。但球体的好处是，绕着任何直径转动都是对称的。

[18]*Brian Malow,Nature 11 December 2008.*

[19]行星并不是直接落在水晶球面上，而是在另一些以球面上的某一点为圆心的圆上。

[20]另一种方法略有不同，也就是让地球略微偏离均轮的圆心。对不同的行星来说，圆心偏离地球的程度也各不相同。如果要进一步提高精度，就要让行星的轨道处于不同的平面中。

[21]这是阿方索十世对托勒密行星系统的数学复杂性的评价，可能是伪托。

[22]在现代天体动力学的研究中，如果一种运动像中世纪学者对托勒密本轮理论的扩充一样，是由数量有限的周期运动叠加而成的，就叫做准周期运动（quasi-periodic）。如果无限多个运动叠加以后，收敛为一个有限的结果的话，那么这种运动就叫做殆周期运动（almost periodic）。

[23] *O. Gingerich,The Book That Nobody Read,* Walker, New York (2004).在《无人读过的书》中，金格里奇对哥白尼《天体运行论》的价值和影响给出了极高的评价。他翻阅了所有现存的《天体运行论》的早期版本，研究它们分别被谁读过以及读得有多深入。

[24]公元前 3世纪，萨摩斯的阿里斯塔克（Aristarchus of Samos，公元前 310年～约公元前 230年）也提出了这个设想。在《数沙者》（*The Sand Reckoner*）一书中，阿基米德（公元前287年～公元前212年）提到了这个设想：“但是，萨摩斯的阿里斯塔克写了一本书，提出了一些假设，最终的结论是，宇宙实际上比我们现在所能想到的情况大很多倍。他假设恒星和太阳都是静止的。地球在绕着太阳转动，而太阳则位于地球轨道的中心。”

[25] G. B. Riccioli,*Almagestum Novum*, Bologna (1651).

第2章
对于自身重要性的执著

对称性让我平静。缺少对称性则让我发狂。

——伊夫·圣·罗兰（1936～2008，法国时尚设计师）

特殊的时间和地点

不要把它想成与赤道平行的纬线，而要意识到当初它也有机会成为赤道。

——马克·吐温（1835～1910）

尼古拉斯·哥白尼的名字已经与一整套世界观紧密联系在了一起。在科学界，人们用“反哥白尼”来作为对人类中心论思想的蔑称。在天文学界，人们常常将人类在宇宙中的位置并不特殊的假设尊称为哥白尼“原理”。我们不该像古代人那样，认为地球是宇宙的中心，而是应该假设宇宙到处都一样，并以此作为我们理论的基础。所以，地球应该是宇宙中一颗典型行星，在一个典型星系中围绕一颗典型恒星转动。

把地球和人类从宇宙中心移走，是科学家们取得的一个重大进步，但如果过于执著于此的话，我们也不得不承认这个原理也有它自身的隐患。尽管我们没有理由相信自己在宇宙中的位置在各个方面都具有特殊性，但否认它在某些方面可能具有特殊性则是过犹不及。

现在我们已经懂得，在宇宙中，生命只能存在于某些特殊地方。显然我们不能活在恒星的中心，那里连原子都不能幸存；也不能活在物质密度过于稀薄的地方，那里形成不了恒星。[1]如果对生命来说，宇宙的这些“典型”环境都过于严酷，使之无法生存和发展，那么我们就不可能处在典型位置上。这种对哥白尼原理的简单节制，对检验现代宇宙学理论而言至关重要。[2]

就像房地产经纪人所说的，地理位置不代表一切。我们也必须考虑到人类在历史长河中的位置。如果宇宙的总体性质随时间变化，例如年龄越大就越热或越冷，那么我们会发现，只有在宇宙的某些特殊时期，才可能允许恒星、行星和生命的存在。今天我们观察到膨胀宇宙的绝大多数重要特征都与这种偏倚（bias）有关联。现在的宇宙看起来已经很老了，因为构建化学复杂性的砖块，如碳、氮和氧等化学元素的原子核，早已在恒星的缓慢核燃烧过程中造了出来，核燃烧过程在超新星爆发时达到顶峰，将这些生命的必需元素在太空中散播开来。①这些成分变成了行星，最终构成了你和我。恒星的这个炼金过程花了数十亿年，所以宇宙才这么老，我们不应感到意外。我们不可能活在一个非常年轻的宇宙里，因为这样的宇宙来不及制造生命复杂性所需的砖块。

① 化学元素都是由更基本的质子、中子和电子构成的。例如，氢一般由一个质子和一个电子构成。碳一般由六个质子、六个中子和六个

电子构成。宇宙诞生后的一段时期内，只有独立的质子、中子和电子，没有任何化学元素，所有元素都是后来形成的。原初核合成只能产生氢、氦和极少量的锂，排在氦和铁之间的元素可以在恒星内的核聚变中形成，排在铁之后的元素可以在超新星爆发时的剧烈核聚变中形成。——译者注

未来肯定会有这样一个时期，最后一颗恒星耗尽了它所有的燃料，然后“死”了，塌缩成一坨致密的遗迹，温度越来越低，或成了一个黑洞。也许这表示说在某个时期之后，宇宙中再也没有任何生命能存活了。一些人认为这是一件非常不能接受的事情，他们相信生命不可能灭绝。[3]显然，我们所知的生命（碳基的和生化的）不可能无限地存活下去。但如果考虑到我们的高科技的未来发展方向，生命或许还有一线生机。持续的微型化可以节约资源，提高效率，减少污染，甚至开发利用量子世界的奇特韧性（flexibility）。宇宙其他地方高度发达的文明可能已经被迫走上了这样的科技发展之路。他们的纳米尺寸的太空探测器、原子级的机械和纳米计算机，就我们对宇宙探测的原始水平是不可能察觉到的。他们输出的废热可能很少，几乎不留痕迹。要想在遥远的未来存活下去，这可能是我们需要采用的演化策略中影响较小的一种。

平等的自然法则

法律面前，人人平等。

——《罗马法》

在哥白尼之后，他的太阳中心说逐渐得到了完善。最后，一个来自英格兰林肯郡的年轻人艾萨克·牛顿（1643～1727），用一个关于运动和引力的新理论对这个图景进行了定量描述。在后来的大约二百五十年中，牛顿的万有引力定律和运动三定律始终是物理学家和工程师认识世界的主要方法。这些定律把先前所有对运动的图像化描述变成了一种精确的数学化描述。其中那些方程（关于变化的“法则”）的解（法则的“结果”）成功地预言了月球和其他行星的运动。其中一个预言就是行星绕太阳的轨道并不像哥白尼假设的那样是圆形，而是椭圆，椭圆有两个焦点，太阳位于其中之一（图2.1）。

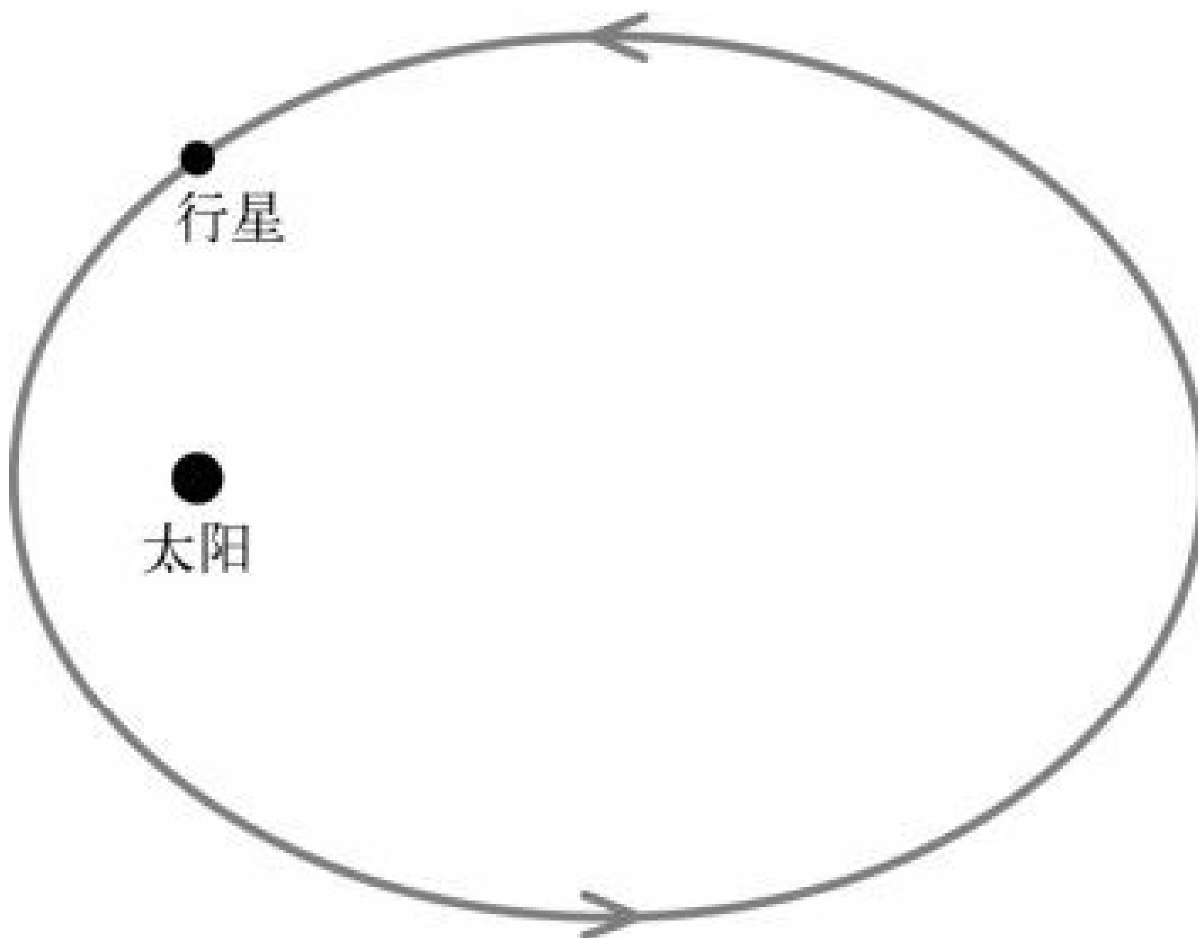


图2.1 一个椭圆的行星轨道，太阳位于椭圆的一个焦点上

牛顿的运动三定律可以被表述成如下形式。

第一定律：不受任何外力作用的物体将保持静止，或做匀速直线运动。

第二定律：物体动量的变化率等于它所受到的外力。

第三定律：每一个作用力都有对应的大小相等、方向相反的反作用力。

这些定律中隐藏着许多值得注意的深刻见解。第一定律提到不受“任何外力作用”的物体。但是谁见过这种物体呢？这是牛顿作为基准的理想化表述。在此之前，大多数人都认为不受外力作用的物体会逐渐减慢并停下来。但牛顿发现减慢是因为有其他外力作用，比如摩擦力和空气阻力。他能敏锐地察觉出在这种情况下所有起作用的外力，并且能够想象出如果所有外力都没有作用其中会怎样，这是一个完全理想化的情况。

当牛顿谈到运动或静止的物体时，我们可能会问：“相对谁静

止？”事实上，他说的所有运动，都相对空间中一个虚构的固定舞台而言，舞台的位置由遥远的恒星划定。他认为恒星是不变的、静止的，这就是后来所谓的“绝对”空间。牛顿定律规定了事物如何在这个舞台里到处活动和表演。人们无论做什么都不能改变这个架构。

牛顿意识到，在宇宙的这个舞台里，只有一些特殊的演员才会认为简单的牛顿运动定律是正确的。这些演员的运动不能有加速度，也不能绕着“绝对”空间里那些遥远恒星转动。假设一位宇航员从转动的太空飞船的窗户向外张望。外面的那些星星在反方向转着圈圈，因而相对他在做加速运动，但这时没有任何外力作用在它们上面。所以，对于自转的宇航员来说，牛顿第一定律失效了，而他导出的第二定律的形式会变得复杂得多。[4]

牛顿的运动三定律反过来揭示了一个问题，既然哥白尼原理存在于自然法则的结果中，那么某种形式的哥白尼原理也应该存在于这些法则本身之中。牛顿定律要求宇宙中存在一类特殊的观测者，对他们来说运动定律比别的观测者看来更加简单。然而，真正的自然法则，如果表述正确的话，应该对所有观测者看起来都一样，无论他们处于什么样的运动状态和位置。没有人有特权认为他们的定律可以比其他人的简单。

物理学家和天文学家借助牛顿定律，就可以试着解释他们看到的所有天体运动。他们可以试着理解星星的分布，理解事物如何从简单的开始变成今天的样子。虽然当时他们缺少我们今天所拥有的望远镜技术，对宇宙景观的视野难免受到技术的限制，但人们还是渐渐开始从天文学的意义上考虑恒星为什么这样分布的问题，并将这些问题与我们所知的物理定律和运动定律联系起来。同时，关键的是，他们开始思考牛顿的运动定律会如何预言宇宙的变化。

变化的宇宙

圆，好似螺旋中的圈/又像转轮中的转轮

——阿兰·伯格曼和玛丽莲·伯格曼（Alan and Marilyn Bergman，美国音乐人）

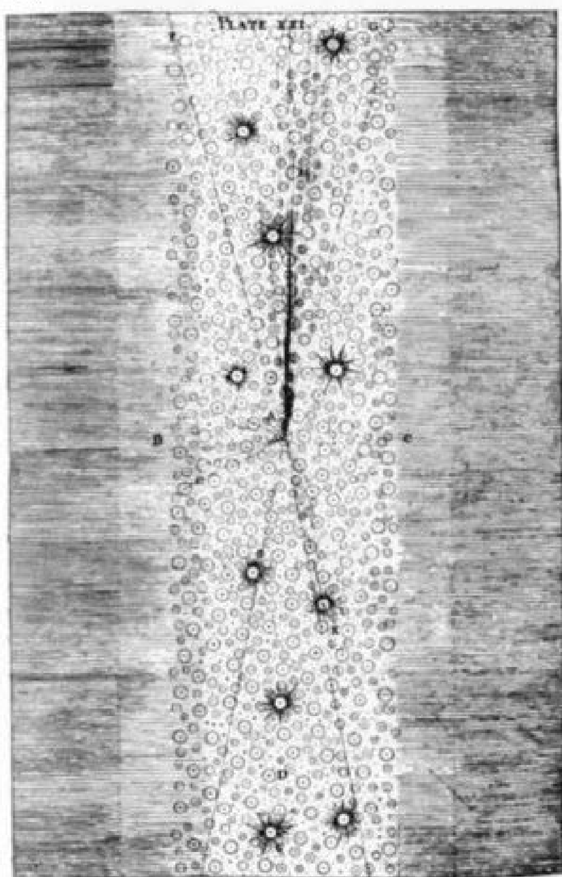
在牛顿之后的几个世纪中，我们对浩瀚宇宙的赞叹之情与日俱增。托马斯·赖特（Thomas Wright，1711～1786）是英格兰北部达勒姆地区的钟表匠，一位自学成才的天文学家、测量员和建筑设计师，他第一次对银河进行了细致描绘。这条由恒星、气体、尘埃构成的闪亮长带自古以来被所有仰望过夜空的人们所熟知和敬仰。[5]借助早期的望远镜，他发现了天空中的星星并不是随机分布的。相反，星星的分布呈现出独特的星系群格局，我们只不过是在其中一个星系的内部向外看。那么，什么样的三维格局才能让银河看起来像现在这样呢？

赖特提出了两个设想。第一个设想是，恒星可能聚集成一个扁平、环形的盘状结构，非常像我们熟悉的土星环，围绕着银河的中心（图2.2a）。银河的中心是“造物中心”，“所有自然法则都源自那里”。第二

个设想是，恒星可能聚集成一个球面，银河只是这个球面上的一块薄片，这体现了我们并不靠近星系中心的事实。

进而赖特的想象又更上了一层楼。他认为宇宙中不应该只存在一个这样的巨大星团。他假设这样的星团应该无穷无尽，每一个都是恒星环绕的中心，有的是球形，有的是盘状。他觉得布满夜空的那些暗淡影像有可能都跟银河一样，它们构成了“无穷无尽的广阔空间……与已知的宇宙并没什么不同”（图2.2b）。

赖特花了大量精力对宇宙进行了充满诗意的图像化描述。他画了一个宇宙的巨大图纸（3米×2米），列出了极为丰富的天文学现象，例如日食和彗星的轨道。他从约翰·弥尔顿的《失乐园》中提到的多个太阳和多个世界的想法中获得了灵感，这种想法认为宇宙由无数个太阳系组成，每一个都有自己的行星，围绕中心的恒星转动。在赖特看来，我们的太阳和别的太阳是一样的，而且地球一定是众多行星当中的普通一员。他估算出银河中大约有超过 3 888 000颗恒星以及“60 000 000个像我们这样的行星世界”，尽管这只是整个夜空中的一小部分。



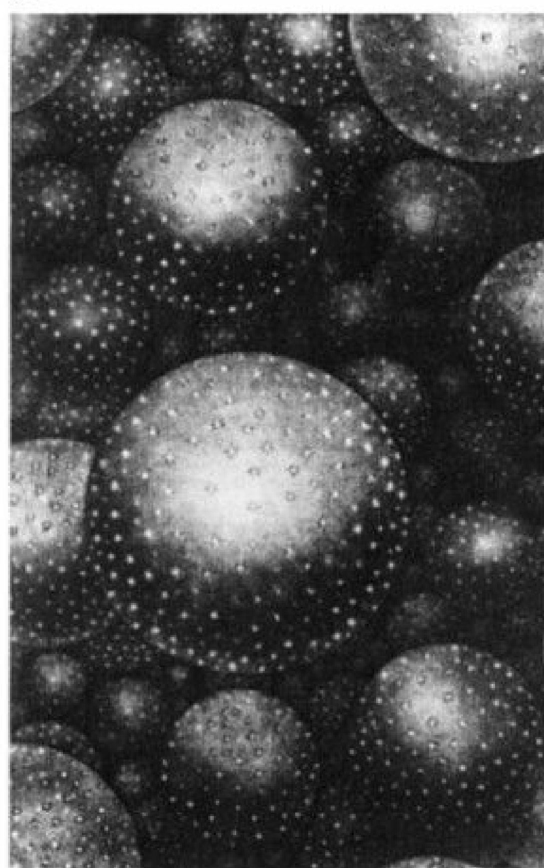


图2.2 (a) 托马斯·赖特的银河系模型，其中的恒星都均匀分布在一个扁平的盘状空间中。(b)赖特的无穷无尽宇宙模型，就像无穷大空间中有无数个泡泡，每个泡泡都是一个星系。两张图都摘自他 1750 年出版的书《一个原创的宇宙理论》（An Original Theory of the Universe）

赖特的设想和他建立的模型将原本着眼于太阳系的哥白尼原理进行了外推，于是就适用于更大的宇宙，这非常重要。宇宙由大量星系（或“岛宇宙”）构成，而我们的银河系只是其中之一。起初，这样的想法并没有被人们认同。到了1920年，华盛顿的史密森尼学会举办了一场著名的大辩论，两位美国天文学家希伯·柯蒂斯（Heber Curtis，1872～1942）与哈罗·沙普利（Harlow Shapley，1885～1972）各执己见。他们争论的焦点是，夜空中看到的螺旋状星云是不是像银河系一样的遥远星系。沙普利声称，银河系就是整个宇宙，但这并不能让当时的天文学家信服。



图2.3 伊曼努尔·康德

具有讽刺意味的是，赖特的高瞻远瞩之所以会被铭记，是因为它以某种方式被后来的科学家所发扬。后来的科学家有了更先进的仪器，于是进一步丰富了赖特的模型。赖特随后将精力转移到了建筑设计领域，并没有为天文学观测作出更多贡献。[6]但是，一位思想敏锐的年轻读者对他提出的宇宙图景十分着迷。1751年，27岁的伊曼努尔·康德（1724～1804）在汉堡市的报纸上读到一篇关于赖特的理论的（不完全可信的）二手说明。四年以后，作为回应，康德匿名写了一本关于宇宙的著作，题为《宇宙发展史概论》。这本书的流传并不广，因为出版商破产

了，印好的册子最后也被执行法官拿走了。不过，仅仅一个世纪之后这本书就声名鹊起，当时德国物理学家赫尔曼·冯·亥姆霍兹（1821～1894）在德国的一次公众演讲时提到了它，遂引起了天文学家的注意。[7]

康德满腔热情地发扬了赖特的银河图景。他提出，赖特的恒星环状结构其实是一个恒星组成的自转盘面，其中向内的万有引力被绕星系中心旋转时产生的向外离心力所抵消。没有引力，这个恒星圆盘就会烟消云散；而没有旋转，所有的恒星就会向星系中心崩塌。康德认为，天文学家在望远镜中看到的所有星云[8]实际上是自转的恒星圆盘。

宇宙中万事万物之间都存在某种相似性，这反映出牛顿引力和运动定律是普适的客观规律。星系之间的差别只体现在视亮度上，也就是反映出它们同我们的距离有远有近；天上的星系千姿百态，又可以解释为它们的朝向与我们视线不平行——就像从不同的角度看一个橄榄球。于是，康德更进了一小步，试图寻找更多星系有可能拥有的形状。如果恒星聚集成像银河系这样的星系，也许星系就会陆续聚集成巨大的星系群，后者再聚集成星系群的星系群，依次类推以至无穷。这个方案并不完全自洽，因为它要求星系围绕星系群的中心旋转，形成盘状结构，而这些星系群之间也都要形成巨大的旋转盘面。不过说到用牛顿定律理解太阳系以外的宇宙结构，这是一次绝妙的尝试。[9]

康德的图景最突出的特点就是宇宙在演化，它就像生生死死的恒星一样会随时间变化。[10]康德的宇宙无边无际，所以并没有真正意义上的中心。但也会有一些特殊的位置，那里的物质密度最大，我们的太阳系就在这样的一个位置。可以看出，生命和有机体像球面波一样，从中心向外传播，所到之处焕发勃勃生机。在每一个尺度上，新形成的结构都在向内的万有引力和向外的离心力作用下保持平衡，就像银河系那样。[11]球面波的前缘布满了其他的太阳系，而其后缘则被耗尽了所有资源、陷入腐朽的死去世界所占据（图2.4）。丰富的新结构在膨胀球面的外部演化着，这是宇宙物质创生的前沿。中心区域的古老腐朽物质是第一批结构形成的残渣。但这不仅仅是个宇宙坟场，这些物质可以重新组合自己，循环利用，将来可以用于产生有序的恒星和行星系统，像是“凤凰涅槃”。因此宇宙可以永远演化下去，“创生永远不会完成或终结，它确实有个开端，但永远不会停止”，并且“各个世界和世界体系毁灭了，被吞进了永恒深渊，但同时创生也总是在天上忙着构建新的结构，好好地弥补那些逝去之物”。[12]

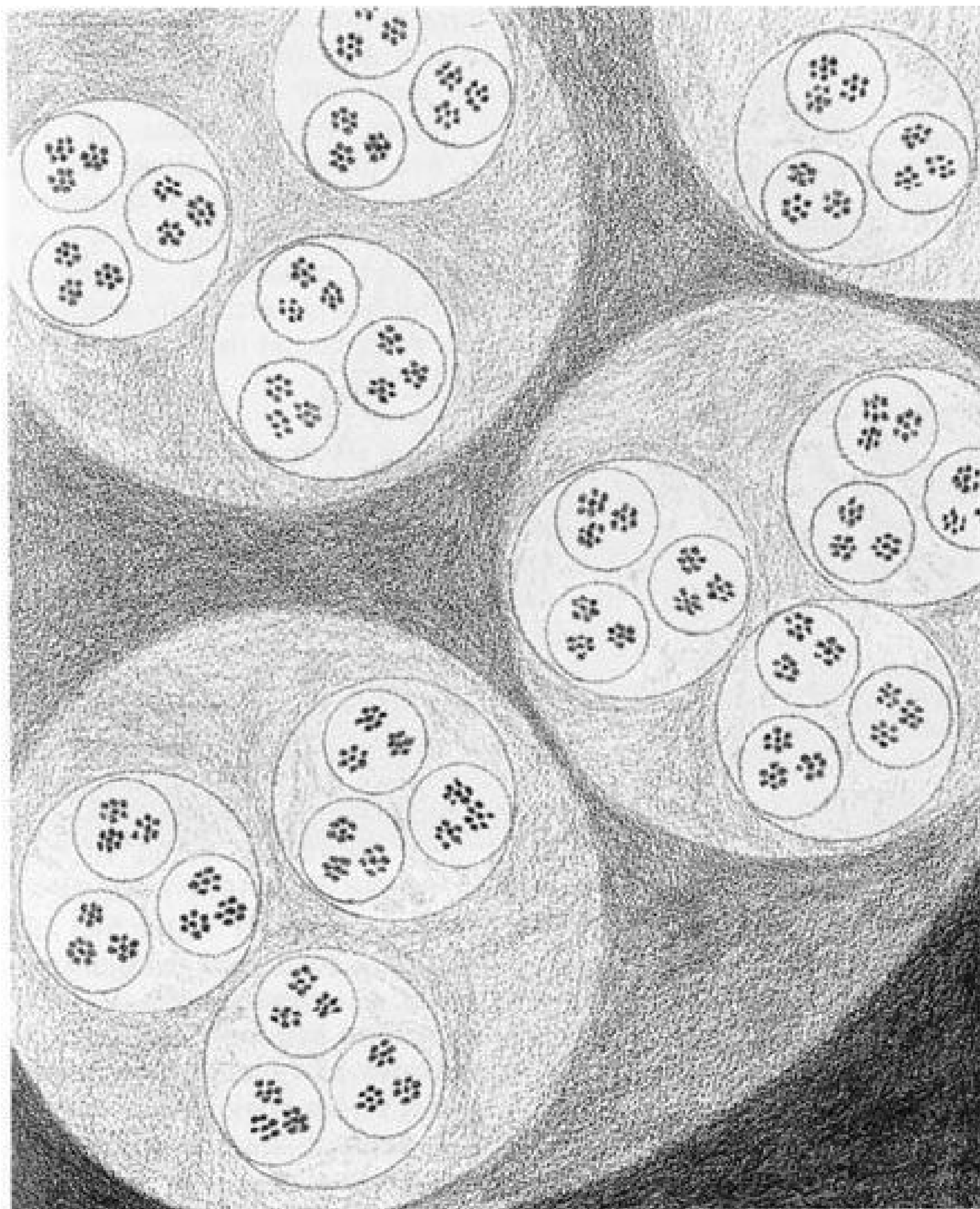


图2.4 康德的演化宇宙模型由无数个像这样的膨胀球面组成，其中每个球面的前缘都能产生恒星，然后渐渐变暗，而其后缘则标志着恒星的消亡[①]

① 作者来信指出，这张图只画出了前缘，而并没有画出后缘。——译者注

康德认为，上帝在不断地创造新的更大的世界，演化的宇宙应该被看作这个神圣伟大计划的一部分。所有已知生命的寿命都是有限的，我们不可能奢望违背这个法则。但大自然是无与伦比地丰富，允许“各个世界和世界体系在扮演过各自的角色之后，退出宇宙的舞台”，同时逐渐探索物质的所有可能组合。

康德把最古老的结构率先腐烂，同时新的结构不断形成的循环叫做“一定的规律”（a certain law）。因此，“已形成的世界被局限在已被消灭的自然之废墟和还未成形的自然之混沌中间”。

在以后的职业生涯中，康德的追求不再是天文学和牛顿物理学，而是批判哲学，这使他成为世界最著名哲学家之一。在1781年撰写的《纯粹理性批判》中，他对实在（reality）和主观实在（perceived reality）进行了重要的划分。我们必须区别客观真理和那些符合我们头脑中特有的范畴而为我们所理解的真理。康德将自己的哲学思辨自比于天文学中的哥白尼革命：①

① 此处参照了蓝公武的译本。——译者注

> 从前，人们假设我们的知识必须与对象保持一致……［但］为了在形而上学上取得进一步的进展，现在我们想试着假设对象必须与我们的知识相一致……这样我们正是沿着哥白尼的思想主线在前进。当假设闪闪发光的漫天星斗在围绕观测者转动时，他就无法对天体运动作出更恰当的解释，于是他把成功的希望寄托在这个假设的反面，即星星是静止的，而观察者在转动。关于对象的直观性，我们也可以在形而上学中进行同样的试验。[13]

对康德来说，事物的绝对真理是不可知的。只有部分，当它符合我们头脑中的范畴时，才可以被我们理解。[14]

星云假说

> 黑暗是永恒的，光明只不过是把黑暗隐藏了起来。
>
> ——丹尼尔·K.麦基南（Daniel Kian McKiernan）

康德对这个宇宙学理论的对立面也作出了贡献。在1755年的《宇宙发展史概论》中，他描绘了另一种图景：太阳系是由自转的气体和尘埃形成的。在1796年的流行读物《宇宙体系论》当中，法国天文学家皮埃尔·拉普拉斯（1749～1827）把这个想法发展成一个更精确的理论。这个关于宇宙本质的想法具有极高的可读性，并对法国（乃至欧洲）的知识界产生了巨大影响。

拉普拉斯可谓法国的大名人，他是拿破仑的科学顾问，杰出的天文学家、数学家和物理学家，并最终被皇帝授予侯爵。他也是个坚定的理性主义者，想要证明不借助任何超自然的力量，也可以解释行星如何形

成。他在那本书的最后一章里解释了太阳系的起源。一团自转的云状物质收缩，形成了一系列行星，在同一个平面上绕太阳公转，同时这些行星也以相同方向围绕各自的对称轴自转。[15]这个理论被称为拉普拉斯的“星云假说”。后来，这个观点在当时的天文学家中很流行，他们认为夜空中每一块发光斑点都是正在形成的行星系统。这与赖特的理论非常不同，后者认为这些发光物应该是像银河系一样的星系，每一个都包含数以亿计的恒星和行星系统。

拉普拉斯侯爵的理论成为了维多利亚时期的宇宙学标准模型。1890年，当时最杰出的天文史学家艾格尼丝·克勒克（Agnes Clerke，1842～1907）曾如此断言：

> 现在完全可以这么说，没有任何合格的思想家……还能坚称任何单个星云都是与银河系相同级别的恒星系统。实际上可以肯定的是，天体中所有的星状和星云状的东西都属于同一个巨大集团。[16]

维多利亚时期的宇宙是由众多恒星构成的大轱辘，也就是我们所见到的银河。“夜空里的点点星光也许是河外星系”的观点暂时运交华盖了。

爱德华时期的宇宙中的生命

> “为什么我们在这儿？因为我们不总是在那儿。”
>
> ——BBC电视剧《探案新窍门》[17]

艾尔弗雷德·罗素·华莱士（Alfred Russel Wallace，1823～1913）是19世纪一位伟大的科学家，他发现生命在自然选择的过程中会发生演化。不过，他并没有为此得到应有的赞誉。达尔文也在一直思考相同的想法，并经过很长一段时间独立地搜集了有关证据。但达尔文幸运的是，华莱士没有简单地先把自己的想法发表出来，而是写信告诉了达尔文。后来华莱士和达尔文的自然选择理论同时公布于世。华莱士对物理学、天文学和地球科学都感兴趣，在很长一段时期内把达尔文当作自己的同事，为了支持他的工作从很远的地方寄来标本。1903年，在《人类在宇宙中的位置》一书中，华莱士广泛研究了地球成为宜居之地的原因以及从宇宙的状态可以得出的哲学结论。[18]

华莱士对开尔文的一个简单的宇宙模型印象非常深刻。开尔文男爵[19]是当时英国科学界的领军人物，皇家科学学会的主席（1890～1895）。这个模型从牛顿引力理论发展而来，可以预测宇宙中巨型云状物质的命运。开尔文的兴趣格外广泛，并且都是从很小的时候就开始的：他从10岁起就在格拉斯哥大学听课，15岁时就写出了关于地球结构的重要研究论文；他发展了能量守恒定律和热力学定律的学说，引入了绝对温标；他也搞过仪器，并于1858年参与了跨太西洋海底电报电缆的研制和调试；他还抽空设计了标准水龙头，设计了中央暖气和空调的热泵，并对第一条电气化铁路的设计作出了重要贡献。

当他转而研究宇宙时，思想的锋芒同样锐利。他证明了，一团球状物质最终会由于自身引力而向中心塌缩，唯一能避免落入中心的方法就是绕着它旋转，而这正是康德提出的设想。开尔文的模型包含了约十亿颗太阳大小的恒星，这些恒星产生的引力导致恒星的公转速度正好与我们在附近的太空中所观察到的符合。[20]

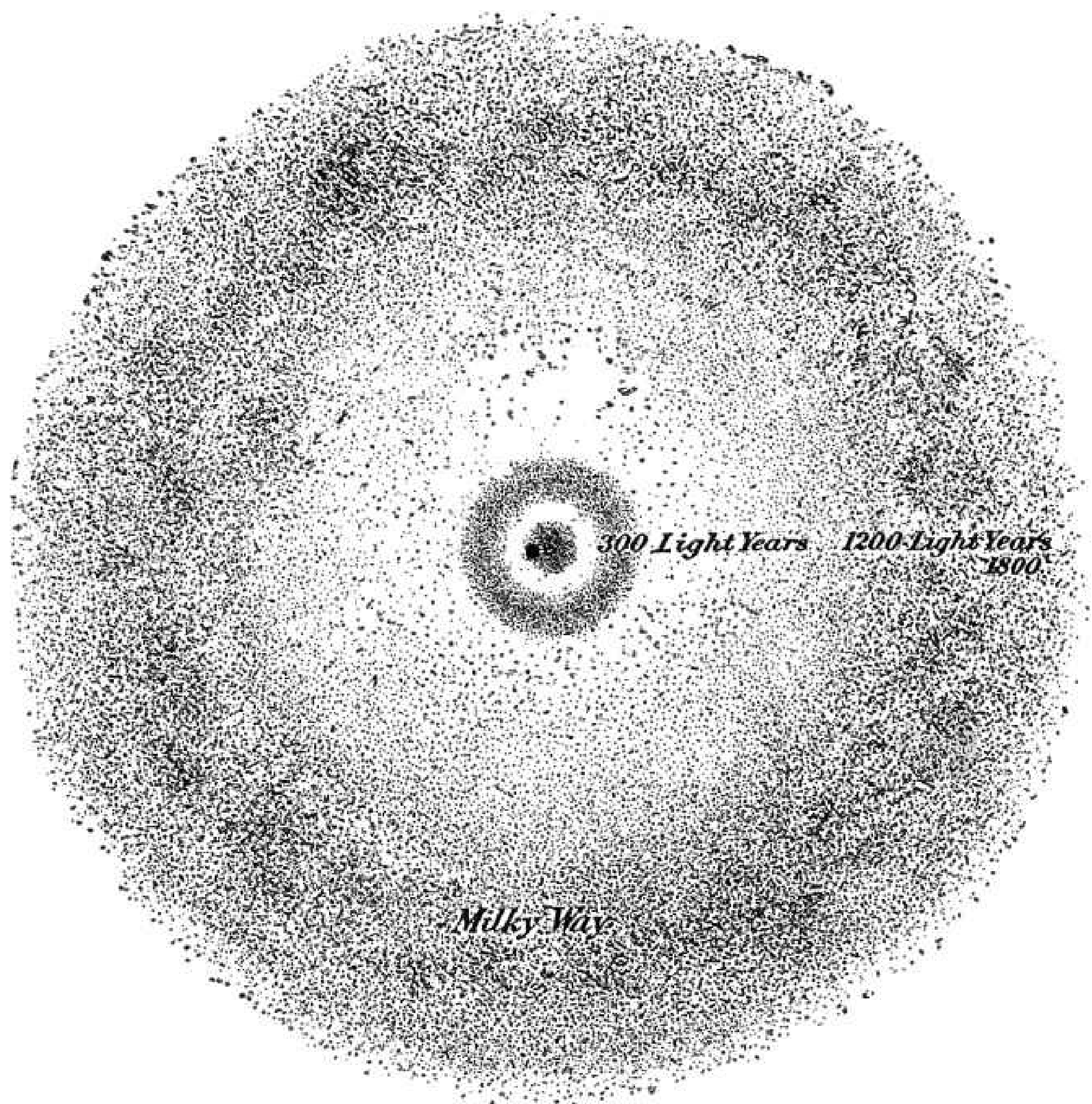


图2.5 华莱士在1903年画的开尔文宇宙模型，其中太阳系（黑点所示）不在中心位置

图 2.5 是华莱士所画的开尔文宇宙模型。[21]华莱士对于开尔文宇宙模型的讨论的有趣之处在于，他采取了一种表面上看起来非哥白尼的观

点，因为他认为宇宙中一些地方比其他地方更有利于生命的出现，并且我们因而并非生活在万物的中心，而是在靠近中心的地方。

在开尔文的宇宙模型里，物质将向着中心掉落，并与已经在那里的恒星融合，从而释放出热量，维持着漫长时间的能量输出。华莱士对宇宙广袤尺寸的解释值得大段摘录：

我们已经发现了，能够合理地解释太阳以及恒星群中其他同等地位的恒星得以长时间发光发热的原因。起初，原始宇宙核心部分弥漫的物质缓慢地运动着，逐渐会聚集成巨大质量的天体。但一段时间之后，会有物质以极高的速度，从周边区域稳步地、均匀地流进来，补充物质以产生和维持像我们的太阳那样的高温，并足以延续生命持续发展所需的漫长时间。对生命这个演化的最终产物来说，原始宇宙弥漫的物质的巨大质量和体积（这由开尔文勋爵假设提出）有着极为重要的意义，因为如果不是这样，中心区域就会相对缓慢而寒冷，不足以持续产生必要的热量；同时，到目前为止占有更大比重的、聚集在银河巨大的旋转环中的大部分物质也同样重要，可以防止有物质过多过快地流入中心区域……对于那些［恒星周边的行星］来说，如果它们的物质演化过快或过慢，就不曾有或不会有足够的时间让生命得以发展。[22]

华莱士看到，宇宙的总体性质和生命出现并欣欣向荣所必需的条件之间存在一定的联系：

我们可以隐约看到恒星宇宙的所有重大特性都与生命的成功发展有关系。这包括，广袤的空间、所形成的银河大圆环的形态，以及我们的位置比较接近它的中心但又不是完全落在中心上。[23]

同时他也发现，物质流入中心，并将引力势能转化为恒星热能的过程是不连贯的，先是很长一段时间的流入，加热了恒星，输出了热量，然后会冷却一段时期，也就是我们正处的时期。

华莱士最后将注意力转移到了地球的地质和历史方面，以讨论生命演化所需的宇宙条件。在那里，他看到的情况远比天文学中的复杂得多。他高度评价了历史的偶然性在人类演化轨迹中所起的主导作用，并认为生命所需的全部条件“绝不可能”在别的地方也找得到。这让他做出了这样的思索：

我们已知的、环绕在周围的广袤而复杂的宇宙，也许是一种绝对的必需……这样才能产生一个所有细节都为生命的有序发展（最终以人类的出现为高潮）做好铺垫的世界。[24]

华莱士厌恶宇宙中还有其他生命的想法，不过他相信物理和化学定律[25]在宇宙中的一致性将保证：

无论有机的生命可能存在于宇宙何处，它们的基础和本质必然是一样的。如果生命存在于某处，它们的外观将会千变万化，就像地球上的生物一样……与我们的宇宙大相径庭的宇宙，其中的物质和我们宇宙中的不同，其中的普遍条件和规律也不同。我们并不是说，在与我们所知所想完全不同的条件下，有机的生命一定不能生存。然而在我们已知的宇宙中，除非满足了我们这里的普遍规律和条件，否则没有理由相信有机的生命是可能存在的。[26]

华莱士的宇宙学研究方法表明，生命演化所需的条件虽然对于研究任何具体的恒星演化理论来说没什么紧密关系，但对于所有宇宙学理论研究来说，它却是必须被适当对待的。

衰败的宇宙

……如果考虑整个宇宙的情况，假如有足够的纸张和墨水的话，我们应该可以写出一个方程，把宇宙的历史往前推，不论多远的未来都可以。但如果我们想往回计算世界过去的历史，就会算到某一个点，在那里方程失去了意义，因为我们会面对事物的这样一种状态，即它

不能在已知的自然法则下，由前一刻的某一状态推导得出。

——威廉·金登·克利福德（1845～1879，英国数学家、哲学家）[27]

19世纪时，一种看待宇宙的新方法（我们称为范式）开始出现。工业革命主导了维多利亚时期。工程、机器、轮船、蒸汽机和熔炉持续推动了经济发展，科学的发展也体现了相应的关切，热力学定律的发现就是集大成者。[28]变化和进步成为了哲学家和工程师的信条之一。所以毫不奇怪，科学家开始将整个宇宙看作一台巨大的机器，并且开始考虑，关于宇宙的过去和未来，热力学定律能告诉我们多少知识。

关于热机，物理学家做出的影响最为深远的发现是，能量会从有序的形式（例如电流或旋转运动）转化成完全无序的形式（例如热辐射）。1850年，鲁道夫·克劳修斯（Rudolf Clausius，1822～1888）证明，在一个任何东西都逃不出去的有限闭合系统中，这种能量的转化是单向进行的。无序的能量，在1865年被他称作“熵”，永远不会减少。这就是所谓的“热力学第二定律”，也是最重要的阐释性科学原理之一。[29]不过在传统的意义上，相比牛顿力学，热力学第二定律还不能算是一个自然法则。它不能回答当一个力作用时会如何，或是一个物体在重力的作用下会如何；热力学第二定律是个统计性定律，它能决定整个系统中所有分子的集体行为。

牛顿定律允许各种各样现实中不存在的事情发生。举个例子，牛顿定律允许葡萄酒杯掉在地上摔成碎片（这个我们能见到），也允许这个过程的时间反演，也就是许多玻璃碎片自然地同时聚集在一起，形成一个完整的葡萄酒杯（这个我们永远也见不到）。导致这个矛盾的原因是，要设法达到玻璃杯破碎的条件并不难，但如果想让所有大小合适的碎片以合适的速度、合适的方向运动，形成一个完整的葡萄酒杯，简直是异想天开，绝不可能。于是，尽管牛顿定律允许这样的现象发生，但我们永远都不会看到一系列由无序转化为有序的事件。相反，我们看到的是由有序向无序的衰败过程，毕竟这更有可能出现。

如果这个无序度增长的“第二定律”适用于整个宇宙会怎样呢？用克劳修斯的话来说，这意味着“世界的熵趋向于一个极大值”。尤其是他认为，热力学第二定律排除了循环宇宙论。在循环宇宙论中，宇宙可以重现一模一样的总体环境，或是像凤凰涅槃一样死而复生。就是这个问题引发了“宇宙热寂说”。从有序向无序的过程是不可逆的，这就意味着，显然宇宙注定会江河日下，持续不断地从有序状态向在遥远未来的无序状态衰败。

最终，世间万物都将淹没在热辐射的海洋里。那时不再有恒星和行星，所有的地方、所有东西的温度和能量都一样。在这个均匀的温度中，不再有任何改变和发展，我们称之为“生命”的现象也灭绝了。追溯过去，这个平稳的衰败过程一定是从过去某个时候开始出现的，那时的

宇宙更加有序。或许宇宙具有一个最大有序度的开端？又或者，应该得到的正确结论是，现在宇宙的年龄不可能是无穷大，不然它早就该达到完全的热平衡和热寂的状态了？[30]

这个想法吸引了那些仍在试图调和两种观点的人，一种观点是宇宙从过去有限的某个时间的“虚空”中产生，另一种则是新兴的变化和进化的观点。尽管对于未来的人类来说，这个想法所描绘的东西太黑暗了。产业革命的进展和技术的变革改变了工业世界的面貌，同时也无情地走向了一个无所谓人类存在不存在的终点。突然间，宇宙看起来不是一个适合生活的地方了。

从1851年到1854年，开尔文在一系列论文和讲座中讨论了这些想法。他感兴趣的是，对于宇宙的过去和未来，第二定律能够告诉我们什么。开尔文有一个强烈的宗教动机，他想推导出一个宇宙的开端，并排除掉永恒循环宇宙论。[31]但是他对热寂的事耿耿于怀，无法接受第二定律这个不可避免的推论。相反，开尔文相信宇宙是无限的（而不像克劳修斯提出的宇宙是有限的），并且相信在未来，自然法则可能是会改变的。别的一些人，例如恩斯特·马赫（Ernst Mach，1838～1916），也试着把热力学定律的推论局限在个别事物上，如恒星和行星，而拒绝把热力学定律应用到整个宇宙。对他们来说，宇宙是一个封闭热力学系统的迹象并不明显，甚至说宇宙受到熵影响的迹象也不明显。

利用第二定律推出宇宙必然有开端的做法并不局限于基督徒，也有热情的唯物主义者。逻辑学家、哲学家、经济学家威廉·杰文斯（William Jevons，1835～1882）[32]，相信第二定律意味着宇宙必然有一个开端，或者说在某个时刻之前，自然法则是不同的。然而政治哲学家如弗里德里希·恩格斯（1820～1895），辩证唯物主义的支持者，却只有在循环宇宙论面前才赞同熵增加原理。而且他认为，所有断定世界有限和世界将在热寂中衰败的观点，都是在暗中承认上帝的存在，因此他要坚决否定。①

① 学者在最近的一篇文章［J. B. Foster and P. Burkett,‘Classical Marxism and the Second Law of Thermodynamics: Marx/Engels, the Heat Death of the Universe Hypothesis, and the Origins of Ecological Economics’,*Organization Environment* 21, 3 (2008)］中指出，从恩格斯的行文来看，他并不是反对热力学第二定律本身，而是反对由此推出的“宇宙热寂说”。——译者注

第一个从数学层面上认真看待这件事的人，是信奉天主教的物理学家和科学史学家皮埃尔·杜亥姆（Pierre Duhem，1861～1916）。他认

为，自然界的熵永远在增加，并不意味着在过去某个有限的时刻它一定是零。[33]他反对用熵增加原理说明宇宙在有限的过去从虚空中产生，或者说明未来将处在一片热寂中，因为宇宙熵的连续增加并不意味着它曾有一个极小值，或者将来必然达到一个极大值。图2.6是一个简单的例子。

从热力学角度看待宇宙的最后一种新设想由路德维希·波尔兹曼（Ludwig Boltzmann，1844～1906）和恩斯特·策梅洛（Ernst Zermelo，1871～1953）在1895年提出。他们认为，宇宙是无限的，而且已经处于整体上的热平衡态，但虽然如此，系统各处都可能存在围绕这个平衡的临时性随机涨落。其中一些涨落可以像银河系一样大，足以成为生命存在的场所。[34]只是这样的大型涨落非常少有，因此生命也非常罕见。

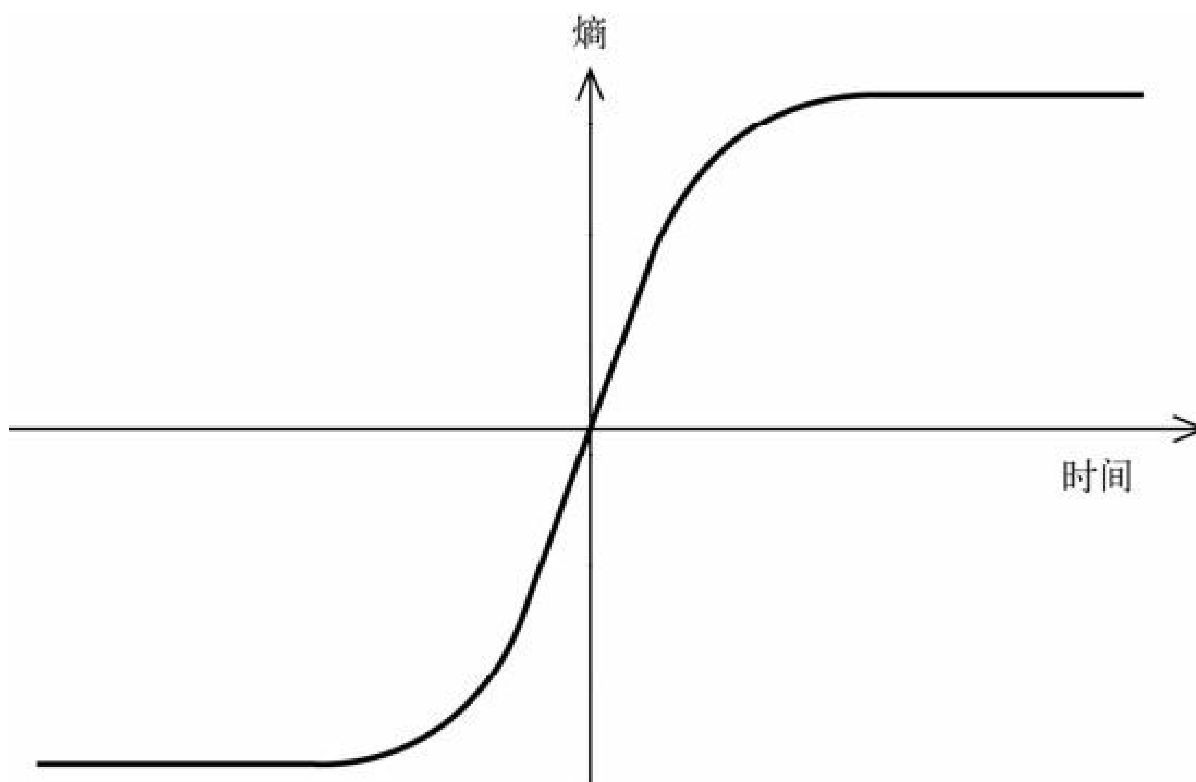


图2.6 一个单调增长的曲线，过去从来不曾等于零，未来的增长也不会超过某个数值

事实上，早在1879年，波尔兹曼的涨落说已经被英国物理学家塞缪尔·托尔弗·普雷斯顿（Samuel Tolver Preston，1844～1917）提出来了。普雷斯顿原本学的是电报工程专业，但后来变成了热力学和引力方面的专家，并最终于1894年（当时他50岁）在德国拿到了他的博士学位。他被宇宙的广袤所震撼，因此觉得我们不能仅从所拥有的一粟中，就推断

出关于沧海的最终结论。他提出宇宙中的一些区域体现出适宜生命居住的特点，但我们不能就因此说整个宇宙都是这样。尤其是我们必定会发现，宇宙中我们所在部分的熵在增加，以便允许生物化学反应的发生，因为“从我们存在于世的事实来看，我们必然位于适宜生命存在的部分”。而且，

> 宇宙在许多广阔区域内都会有这样的奇特性质，允许温度、聚集状态、组成以及形成宇宙的物质发生几乎是随机的局域涨落……整体的构成（宏观地看）仍然是彻头彻尾一致的。[35]

对于这个理论，有人批评，如果各处的熵都以相同的速率增加，那么就需要假设已知的物理定律在过去的某时失效了。对此，普雷斯顿成功地进行了辩驳。[36]我们将在第10章继续讨论普雷斯顿的这个想法，在被提出一百三十多年之后，它对宇宙学来说仍然是有意义的。

卡尔·史瓦西：超越时代的人

> 我已经受够了数/天上的星星
>
> ——艾弗利兄弟二重唱[37]

19世纪时，数学家们终于开始了解这么多个世纪来一直摆在他们面前的事实了。此前，除了欧几里得几何对平面上的线、点和角的经典描述，他们并不接受其他几何系统的存在。欧式几何被看作其领域内唯一正确的逻辑系统，这个偏见深深地植根于一个信念之中，那就是它和宇宙有相通之处。这并不仅仅是个数学“游戏”，从一系列初始假设和规则开始，就可以得出所有可能的几何结论。这是世界的真实存在：一条关于事物本质的绝对真理。当追问上帝或宇宙终极本质的神学家、科学家或哲学家被人批评，被问到其中可有什么东西已经完全搞清楚了时，他们就会以欧式几何为例，说明人类的思想如何把握住了一条终极真理。这就是为什么有时候他们会把自己的论文仿照欧几里得的方式进行论述。欧几里得就是黄金规则。

曲面上也可能有逻辑自洽的几何学，例如马鞍形面或球面，这个发现已经被航海家和艺术家直观地使用了数百年，他们对此并不感到惊奇。然而这些曲面上的几何却意外地引发了人类的思想革命。突然间有了许多种可能的几何。每一种都不过是一种从初始假设和规则推导得出的逻辑体系。没有一种几何敢于断言自己是终极真理的一部分。于是，几何和所有的数学都换了个态度，认为自己是一种公理和规则的体系。这些几何在逻辑上都自洽，在这个意义上它们都是“存在的”，但这并不能让它们成为真实的或必然的物质实在。

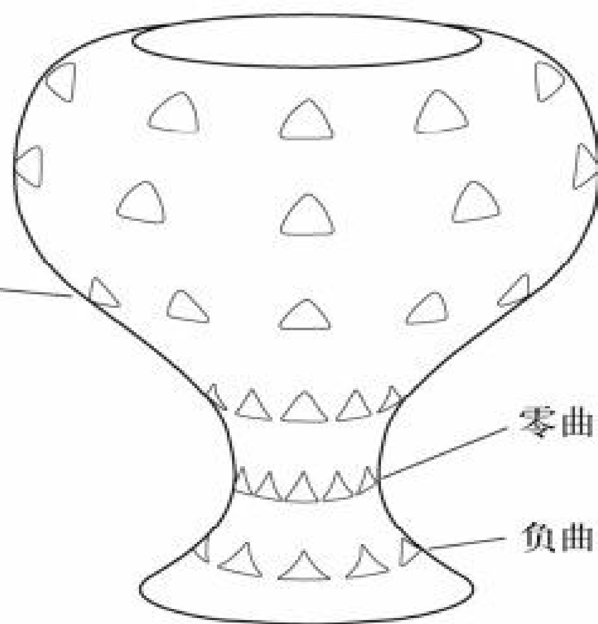
非欧几何最简单的例子是那些描述负曲率或正曲率曲面的几何，就像这里你看到的花瓶的图案（图 2.7）。一个曲面可以非常复杂，有些部分的曲率可以是正的、负的，也可以是零。想要了解曲面的曲率，最简单的办法是选取其中三个相邻的点，A、B和C，画出能够连接A到B、B到C、C再回到A的最短曲线。在一个平面上，这些最短的曲线其

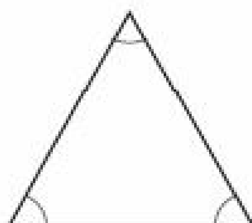
实就是直线，ABC就是一个内角和为180度的三角形。

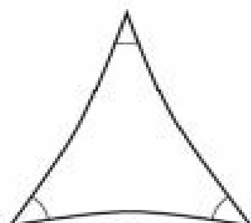
在一个正曲率的面上，比如球面，连接 A、B和 C的最短距离就不再是像平面中那样的“直线”了。这些曲线其实是绕球心的圆弧。它们就是“大圆”，洲际航班为了将油耗降到最低而采用的路线（假设风都很小）。它们闭合后形成一个内角和大于180度的凸三角形。这是正曲率的特征。类似地，在一个负曲率的面上，比如一个马鞍面、一块品客薯片、一片冬青的叶片或是羽衣甘蓝的叶片（图 2.7b）[38]，其中的弯曲三角形的内角和小于180度。

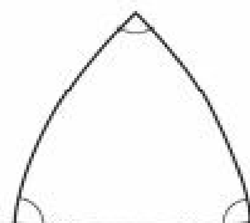
有时候所谓的弯曲并不与我们朴素的直觉精确相符。从外观上看，你也许认为圆柱体是弯曲的，但事实并非如此。如果你拿一张平展的长方形纸，在上面画一个三角形，那么其内角和正如预料的那样是180度。现在把这张纸的两个长边黏在一起，形成一个圆柱体，让三角形朝外。三角形看起来跟往常一样，内角和是 180度。局部地看，圆柱体表面并不是弯曲的（图2.8）。

(a)









(b)



图2.7 (a) 连接弯曲表面上间隔不远的三个点所形成的三角形。花瓶底座的表面是正曲率的，于是三角形的内角和大于180度；瓶口附近是负曲率的表面，于是三角形内角和小于180度；这两个区域之间是曲率为零的表面，三角形内角和等于180度。(b) 甘蓝是卷心菜的一个变种，叶子表面是负曲率的

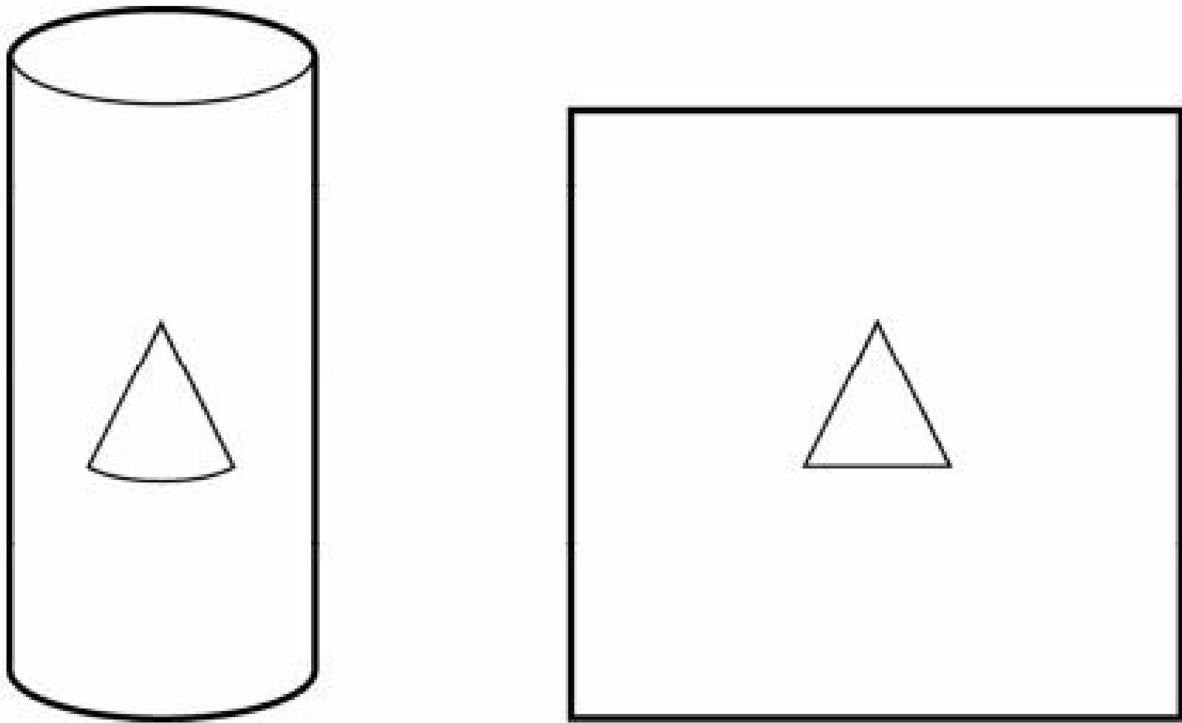


图2.8 黏合正方形的两个对边就形成了一个圆柱体。从局部看，圆柱体的几何仍然是平坦的，跟正方形一样。圆柱表面的三角形内角和仍然是180度，也跟正方形一样

卡尔·史瓦西（Karl Schwarzschild，1873～1916）是一个天才，可惜英年早逝，无法亲眼目睹自己想法的真正重要性。1916年5月，他年仅42岁就去世了。在恒星、星系和引力的研究中，他作出了许多发现。他发现了如今遍布宇宙的黑洞的精确解，为爱因斯坦革命性的相对论的精确实验检验铺平了道路。但在此之前，在1900年，从弯曲几何的理论中，他提出了一个宇宙的新图景。1900年7月，在海德堡召开的德国天文学会的会议上，史瓦西提出，宇宙的几何性质并不像欧几里得教给我们的那样是平坦的，而可能是弯曲的非欧几何。非欧几何在18世纪早期由约翰·兰伯特和意大利耶稣会士、数学家乔瓦尼·萨凯里（Giovanni Saccheri，1667～1733）首先提出，并由黎曼、高斯、波尔约和罗巴切

夫斯基[39]于19世纪早期进一步发展。[40]这些新的可能性并没有受到物理学家和天文学家的一致欢迎，就连最富远见卓识的物理学家如詹姆斯·克拉克·麦克斯韦（1831～1879），都在1874年给他的苏格兰老朋友[41]彼得·泰特（Peter Tait，1831～1901）的一张明信片中，把支持这种几何扩展的人称为“空间折皱者”（space crumplers）。[42]

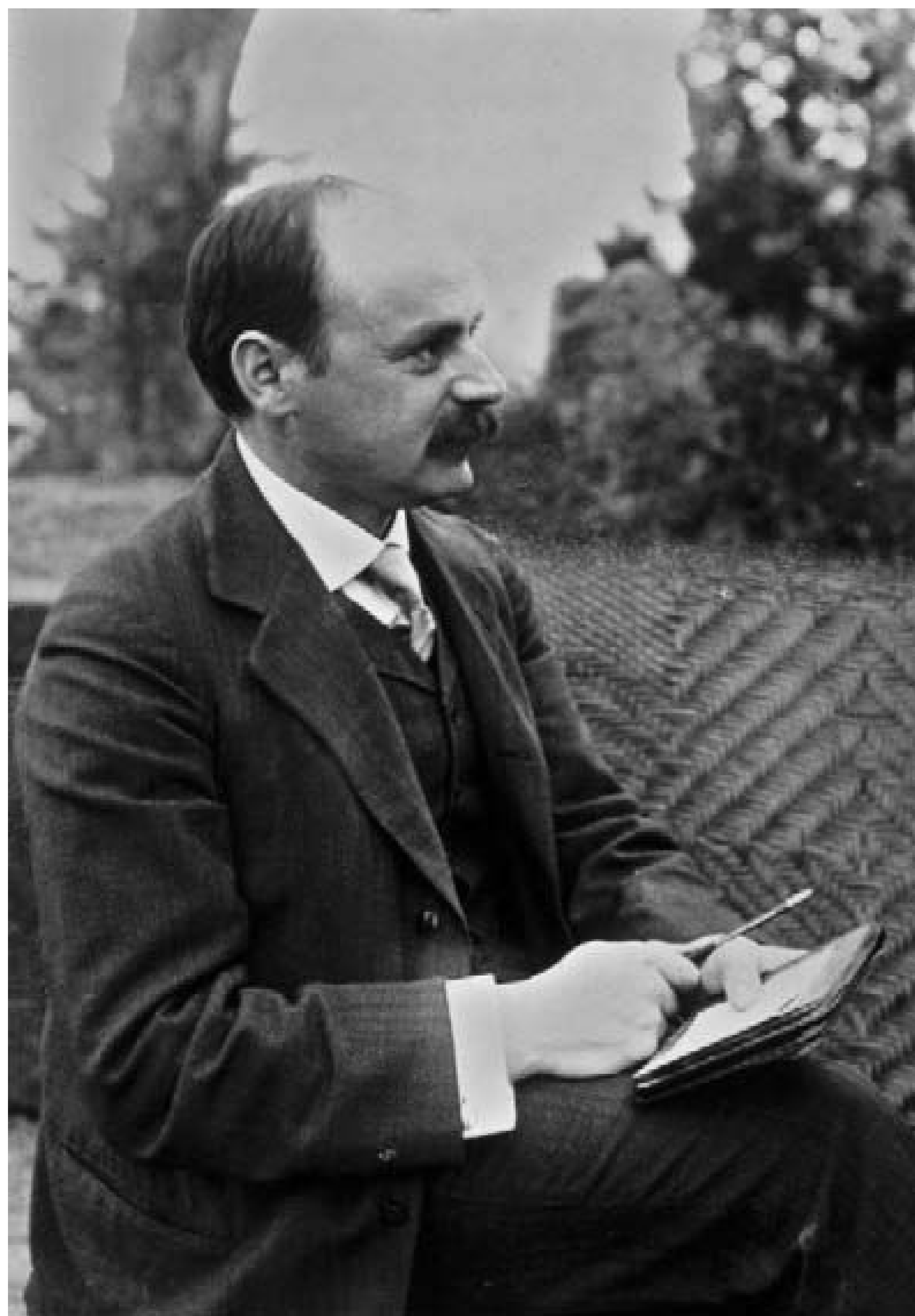


图2.9 卡尔·史瓦西

史瓦西首先意识到，如果宇宙有负的曲率，那么恒星的视差角①就会有一个极小值，正如罗巴切夫斯基所指出的，于是他推导得出宇宙的曲率半径必然大于60光年。更有趣的是，他接着又考虑了宇宙有正曲率的情况。这意味着宇宙是有限而无界的，就像一个球面，自我闭合。[43]

① 从两个不同的地点看同一个物体，视线的夹角就是视差角。天文学中常用地球公转产生的视差角来测量恒星到太阳的距离。——译者注

史瓦西发现，100颗由视差法已测得距离的恒星以及 1亿颗由于视差角太小（小于 0.1 角秒）而测不出距离的恒星，可以很好地安放在半径不小于2500光年的正曲率球形空间里而不显得拥挤。他又提出在这样的空间里，如果我们朝太阳的反方向看，原则上[44]可以“看到”另一个太阳的像，因为光线绕着这个球形空间跑了一圈。

第一次世界大战爆发后，卡尔·史瓦西志愿加入德国军队。在俄国前线时，他写下了两篇绝世之作，一篇关于量子理论[45]，一篇关于爱因斯坦的相对论，两个都是诺贝尔奖级别的。1916年，他得了天疱疮，那是一种由免疫系统崩溃而引起的严重皮肤病，当时还没有治疗方法。到了3月，他被带回了家中，仅仅2个月后就去世了。

旧的世界观至此到了终结之时。19世纪的宇宙观曾以标新立异为耻，直到非常晚的时期，这种情况才有所改观。史瓦西的想法在当时并没有引起多少人注意。至此所有的牌洗得只剩下两张，一张是宇宙由许多星系组成，一张是只有一个银河系，其他所有的遥远星云都位于其中，虽然古代人可能认为宇宙的结构还存在别的可能性。但接下来，人类思想所能触及的领域将会大幅度地向外拓展。

注释

[1] 这个总体的思路叫做人择原理，是布兰登·卡特（Brandon Carter）于1973年在纪念哥白尼诞辰500周年的一次天文学演讲中提到的。进一步的讨论参见：J. D. Barrow and F. J. Tipler,*The Anthropic Cosmological Principle*, Oxford UP, Oxford (1986).更多应用人择原理的最新例子，参见：J. D. Barrow,*The Constants of Nature*, Jonathan Cape, London (2003).

[2] 一个理论也许会预言说，宇宙可能拥有许多不同的性质，每种性质出现的概率各不相同，而宇宙被观测到的性质或许是其中出现概率很小的结果。但是我们并不能因为这个理由就否定这个理论。也许，只

有在那些出现概率很小的宇宙中才会允许观测者的存在。这种想法很重要，它能够用来检验本质上包含随机性的宇宙学理论，根据这些理论会得出许多性质不同的宇宙。在任何量子宇宙学理论中，这种随机性都是不可避免的。

[3] 物理学家保罗·狄拉克强烈赞成这种观点。

[4] 往牛顿方程中加入的额外项叫做科里奥利力，这是由科里奥利（Gaspard-GustaveCoriolis，1792～1843）于 1835年提出的。在一个旋转的参照系中，运动的物体会受到一种额外的加速作用，使得运动方向发生偏转，这就叫做“科里奥利效应”。同年，科里奥利利用牛顿力学原理，写下了一本关于台球研究的权威著作，叫做《关于台球中的自转、摩擦和碰撞的数学理论》（*The Mathematical Theory of Spin, Friction and Collision in the Game of Billiards*）。

[5] 在希腊语中，银河就是galaxias，由此才有了英语中的galaxy（星系）。

[6] 到了20世纪60年代，人们才发现赖特还有一份未曾发表的草稿，叫做《关于宇宙理论的反思或奇思》（‘Second or Singular Thoughts upon the Theory of the Universe’）。在这篇文章中，赖特放弃了他先前的高瞻远瞩，而换成了一种由无穷多层同心球面组成的宇宙，恒星坐落在这些球面上，球面的共同中心就是上帝。灵魂穿梭于不同的球面之间，分别受到或严格或宽松的限制，这就是神的不同程度的惩罚措施。这种模型的动机来源于当时的炼金术对于火以及神在创造太阳时的所作所为的认识。火的循环维持了宇宙的统一，而彗星正是生命之火在宇宙中扩散并为太阳补给燃料的手段。更详细的讨论可以参见：Simon Schaffer, *J. Hist. Astronomy* 9, 180–200 (1978).

[7] 关于这次演讲可参见最近的一本亥姆霍兹文集：H. Helmholtz,*Science and Culture: Popular and Philosophical Essays*, ed. D. Cahan, University of Chicago Press, Chicago (1995), p. 18.

[8] 在20世纪中叶以前，“星云”一词都被天文学家们用来描述所看到的遥远天体发出的模糊光斑，其中既有恒星又有星系。例如，埃德温·哈勃就把星系称作星云。但在今天，天文学家仍然保留了“星云”的叫法，用来描述恒星周围充满了尘埃和气体的区域。尘埃和气体中的原子和分子与恒星发出的辐射相互作用，能够发出各种颜色的光。不同的情况有不同的叫法，如行星状星云（这个名字跟行星没有任何关系）、发射星云和反射星云等。人们最常拍摄的就是这些壮观天体，它们的照片常见于天文学杂志和海报。

[9] 不久以后，瑞士数学家约翰·兰伯特（Johann Lambert，1728～1777）提出了一个类似的宇宙模型，认为恒星会聚集成不同尺寸的星系

团，然后星系团再聚集成更高一级的星系团，依此类推。这和康德的无限大宇宙不同，兰伯特的模型虽然很大，却是有限的，而且是周期性的，其中每一级结团的图案都是一样的，像是一种分形结构。

[10]更多技术性讨论参见：Barrow and Tipler,*The Anthropic Cosmological Principle*, p. 620.

[11]如果一个系统在距中心为r的地方，物质的旋转速度为v，而且密度随半径的分布为ρ(r)的话，那么v^2一定正比于$\rho(r)r^2$。

[12] I. Kant,*Universal Natural History and Theory of the Heavens*, transl. W. Hastie, University of Michigan Press, Ann Arbor (1969), p. 149.在这个版本中，康德在 1751年对赖特的理论（他正是受其启发）所作的评论也被翻译成了英语。

[13]转引自：D. Danielson (ed),*The Book of the Cosmos: A Helix Anthology*, Perseus, New York (2000), p. 271.

[14]如果康德当初知道自然选择会导致生物演化的话，可能就不会表述得那么极端了。我们的感官是演化的结果，是人类对“实在”的适应，而不论我们对“实在”是否有了解。例如，我们眼睛的结构能够感受光的真实特性，这就是演化的结果。利用光学理论，我们能够直接而精确地理解眼睛的结构，这一事实反过来说明，我们的理论抓住了我们视觉所需的光的大部分本质。参见：J. D. Barrow,*The Artful Universe Expanded*, Oxford University Press, Oxford (2005), pp. 30–33.

[15]当时的人们还不知道天王星的其中两个卫星的运动方式，威廉·赫歇尔（1738～1822）在1798年才发现了它们。（天王星某些卫星的转动方向和太阳系绝大部分天体是相反的，也就是在逆行。——译者注）

[16] A. Clerke, *The System of the Stars*, Longmans, Green, London (1890).转引自：Edward Harrison,*Cosmology*, 2nd edn, Cambridge University Press, Cambridge (2000), pp. 77.

[17]在2008年4月4日BBC1播出的《探案新窍门》中，饰演布莱恩·莱恩的阿伦·阿姆斯特朗（Alun Armstrong）如是说。

[18] A. R. *Wallace*,*Man’s Place in the Universe*, Chapman and Hall, London (1903).此处引用的是1912年出版的该书第四版。

[19]开尔文原名威廉·汤姆逊（William Thomson，1824～1907），于 1892年被授予开尔文男爵一世的贵族身份。他去世后被安葬在威斯敏斯特教堂。

[20]他认为，如果宇宙中存在100亿颗恒星，它们的运动速度就太快了。如果一个引力系统的总质量为 M、半径为 R、平均速度为 v，那么这三个物理量的关系就是v^2 2G M R ，其中G是牛顿引力常数。

[21] Wallace,*Man’s Place in the Universe*, p. 248.

[22]同上，pp. 255, 261.

[23]同上，p. 256.

[24]同上，pp. 256–257.

[25]通过观测木星卫星的掩食现象而测出的光速竟然和地球实验的结果相吻合，这令他非常震惊。于是他认为："众多科学发现让我们坚信，无论是从物理定律和化学定律的作用来看，还是从宇宙的结构和表现形式的力学关系来看，物质宇宙的各个部分本质上是统一的。"（同上，p. 154）

[26]同上，pp. 154–155.

[27] W. K. Clifford,Lectures and *Essays*, vol. 1, Macmillan, London (1879), pp. 221.

[28] S. Brush,*The Kind of Motion We Call Heat*, vols. 1 & 2, N. Holland, Amsterdam (1976).

[29]热力学第一定律说的是能量守恒。

[30] J. Vogt,Die Kraft, Haupt & Tischler, Leipzig (1878), and H. Kragh,*Matter and Spirit in the Universe; Scientific and Religious Preludes to Modern Cosmology*, Imperial College Press, London (2004).

[31]这种动机也可以在更近一些的历史中读到，例如：S. Jaki,*Science and Creation*, Scottish Academic Press, Edinburgh (1974).

[32] W. Jevons,*The Principles of Science*, 2nd edn (1877).

[33]注意，如果一个量在不断变大（例如熵），并不能说明它曾经等于零。

[34] L. Boltzmann, *Nature* 51, 413 (1895).更深入的讨论参见：Barrow and Tipler, *The Anthropic Cosmological Principle*, pp. 173–178.

[35] S. T. Preston, 'On the Possibility of Explaining the Continuance of Life in the Universe Consistent with the Tendency to Temperature-Equilibrium',*Nature* 19, 462 (1879).

[36]威廉·缪尔（W. Muir）给《自然》杂志写的信以及普雷斯顿的回应，参见：'Mr Prestonon General Temperature-Equilibrium',*Nature* 20, 6 (1894); 'Temperature Equilibrium in the Universe in Relation to the Kinetic Theory',*Nature* 20, 28 (1894).

[37]取自歌曲《Bye Bye Love》，由布莱恩特夫妇（Felice and Boudleaux Bryant）作词，艾弗利兄弟在1957年录制。

[38]蔬菜的叶子、花和珊瑚的表面常常是负曲率的，因为这样表面积更大，能够提高营养的吸收效率。

[39]约翰·兰伯特在 1766年首次提出了负曲率空间的想法，高斯（1777～1855）在 1816年设想了如何通过实验来测量地球表面的曲率

大小。1829 年，尼古拉·罗巴切夫斯基（Nicolai Lobachevskii，1792～1856）写了一本书，叫做《几何学原理》（*The Principles of Geometry*）。在这本书中，他证明，欧几里得著名的“第五公设”（平行线永远不会相交）独立于其他的几何公理，如果这个公设不成立的话，就会产生新的几何。匈牙利数学家雅诺斯·波尔约（János Bolyai，1802～1860）也作出了类似的发现。波恩哈德·黎曼（Bernhard Riemann，1826～1866）后来将这类问题写入了他的博士论文，将它发展成了一门关于弯曲空间（或黎曼空间）的完备学科，论文的导师之一正是高斯。

[40]其他人也想知道用弯曲的几何来描述宇宙空间是否更合理。西蒙·纽康（ Simon Newcomb，1835～1909）在《超空间的哲学》［‘The Philosophy of Hyperspace’,Bull. Amer. Math. Soc. (2) 4, 187 (1898)］一文中，介绍了有限的球形空间的优点。关于非欧几何和天文学的早期研究，更完整的介绍参见：D. M. Y. Sommerville, *Bibliography of Non-Euclidean Geometry*, University of St Andrews and Harrisons & Sons, London (1911).

[41]麦克斯韦和泰特曾共同就读于爱丁堡学院。在他们的学术生涯中，他们在数学和物理学领域角逐各种荣誉和奖项。1852年，时年 20 岁的泰特被认为是剑桥大学历史上最年轻的数学荣誉学位考试第一名（senior wrangler）。第二年，麦克斯韦也获得了这项殊荣。泰特最著名的工作是创立了数学中的纽结理论，而麦克斯韦是仅次于牛顿和爱因斯坦的最伟大的理论物理学家。

[42] P. H. Harman,*Energy, Force and Matter*, Cambridge University Press, Cambridge (1982).

[43]事实上，这意味着宇宙是一个假想的四维超球体的三维表面。

[44]实际上，星系间的物质会让光线变模糊，结果就看不到了。

[45]他证明人们观测到的斯塔克效应（Stark effect）可以用量子理论来解释，并发现了爱因斯坦广义相对论中最重要的解析解。他的解最初描述了像太阳那样的球对称物体所产生的引力场，后来略作修改后可用来描述静止不旋转的黑洞所产生的引力场。“史瓦西解”是如今世界上任何地方任何一门引力或天文学课程都要讲到的核心内容。

第3章
爱因斯坦的宇宙

爱因斯坦每天都给我解释他的理论，当我们抵达港口时，我确信他理解了这套理论。

——柴姆·韦兹曼（Chaim Weizmann，1921年和爱因斯坦一同乘船跨越大西洋前往纽约）[1]

完成哥白尼世界观

政府律师文化中的一个悖论让我很震惊。法律条文越不确定，他们的观点反而变得越确定。

——杰克·斯特劳（Jack Straw，英国前司法大臣）[2]

19世纪天文学家描绘的所有宇宙图景都应用了牛顿在1687年提出的技术指南。他著名的运动定律和引力定律对于各种实用场景都非常有用，不论是建造桥梁、测试汽车、导航飞机，还是投掷石块。但是如果你细想一下，就会发现这其中也隐藏着深刻的问题：可惜这些定律只对一类特殊的观测者成立——那些没有相对于无穷远处的恒星转动或做加速运动的观测者。正如我们在上一章中谈到的，如果从旋转的火箭窗口向外看（图3.1），你就会发现尽管星星没有受到任何作用力，但也在做加速运动。



图3.1 从不断旋转的火箭内眺望窗外，宇航员发现恒星都在做加速运动，尽管没有任何力作用在恒星上面

爱因斯坦认为，这种表述自然法则的方式存在严重的问题。某种自然法则的表述只是针对某类观测者，这样在这些观测者看来，世界一下子变得如此简单。这就等于说，世界上存在一种特殊的知识，只有某些人才能搞懂，而别的人都不懂。在爱因斯坦看来，这样的逻辑难以接受，它在本质上是一种前哥白尼世界观——在宇宙中，我们的运动（而不是位置）被赋予了一种特殊的地位。

爱因斯坦的伟大之处在于，他找到了一种方法，使我们在寻找和表述自然法则时，能保证从所有的观测者身上都能得出相同的定律，而不论他们如何运动。爱因斯坦新的引力定律（广义相对论）取代了牛顿的万有引力定律，它也被许多人认为是人类思想最伟大的创造之一。广义

相对论扩展了哥白尼的观点，它不仅仅要求我们在宇宙中没有特殊位置，还要求自然法则本身对所有的物理学家都是一样的，不论他们身在何处，也不论他们如何运动。在实际情况中，这又意味着什么呢？

设想你朝天上看时发现，当A发生的时候，总是导致结果B。因此你发现了一个自然法则，我们记作方程 A=B。然后你把所有的测量设备都放进宇宙飞船里，随飞船进入太空。你的宇宙飞船可能在以一种非常复杂的方式转动或做加速运动。当你仔细观察A时，它看起来总是不太一样，不像地面上测到的结果。由于你的运动影响了观测结果，所以你测到的不再是 A，而是 A'。同理，另一个测量结果不再是 B，而应该是B'，其中B'就是考虑了你的运动产生的影响之后得到的B的大小。爱因斯坦的理论能够保证，无论你的飞船在做什么运动，自然法则都会满足A'=B'。虽然这些A和B有不同的取值，但无论是在飞船上还是地面上，定律的形式都是一样的，如A=B，或A'=B'。牛顿定律并没有这个特点。在牛顿定律中，如果在一个不旋转的飞船中得到A=B的定律，针对同样的事件，在另一个旋转飞船中得到的定律就会变得复杂许多：

$$A'=B'+\text{额外项}$$

在爱因斯坦看来，一个完整的哥白尼世界观不仅要适用于自然法则的结果，如行星、恒星和星系，而且也必须适用于自然法则本身。

爱因斯坦的洞见

> 宇宙不得不被加以再造，而爱因斯坦的原理就像一盏小灯照亮了他的道路。他描摹出一个引力理论的轮廓，在其中，宇宙中所有的物质、所有的粒子、所有的能量都对它的结构有影响：时空被物质和能量压弯了。
>
> ——让·艾森施塔特（法国物理学家）[3]

1931年，爱因斯坦在美国同主演查理·卓别林一起参加了无声电影《城市之光》的首映式。人群中不断传来的喝彩声向这两位不同领域的名

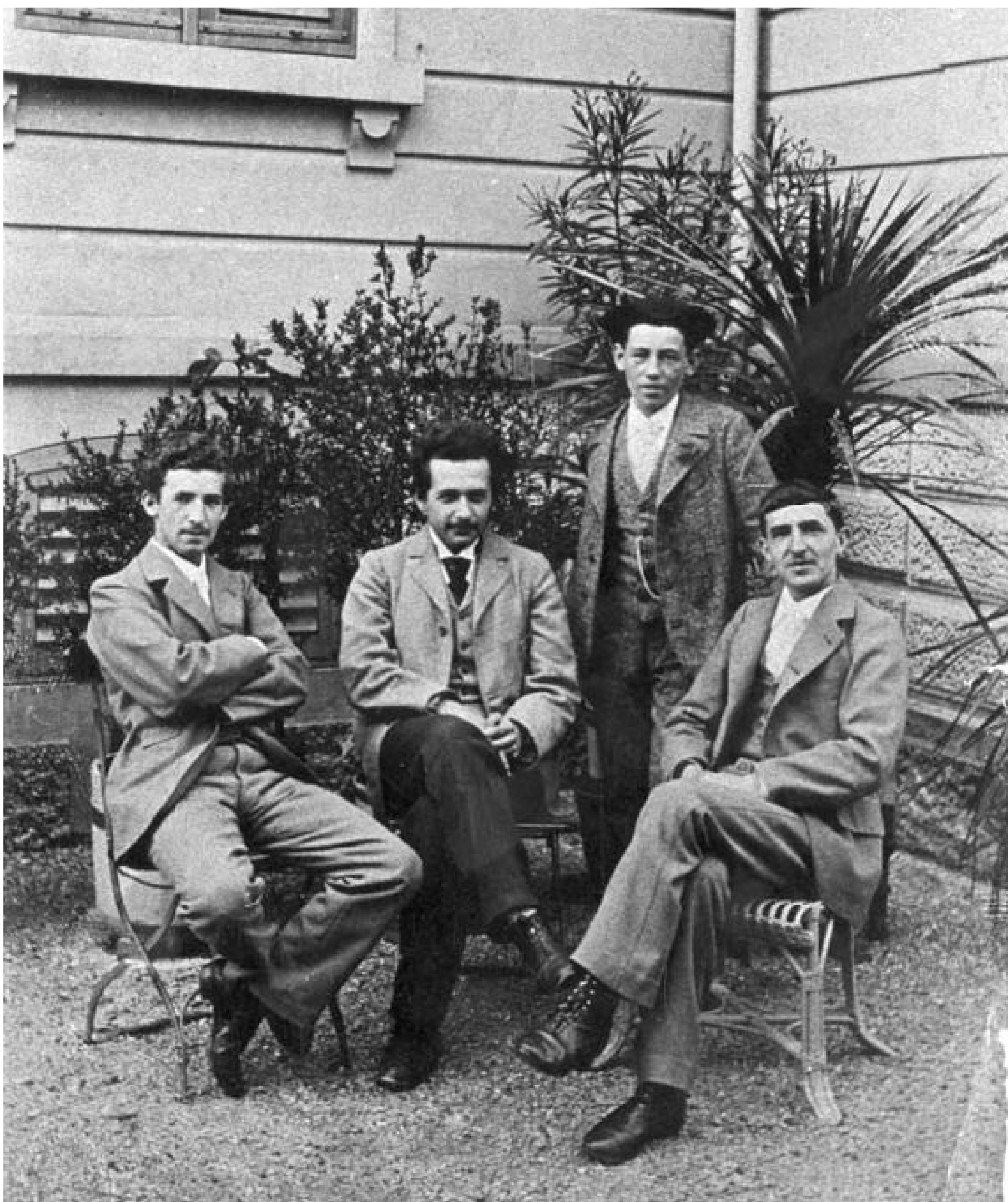


图3.2 在苏黎世求学时的（左起）马赛尔·格罗斯曼、爱因斯坦、古斯塔夫·盖斯勒（Gustav Geissler）和尤金·格罗斯曼（Eugen Grossmann）

人致意。据说，卓别林对爱因斯坦这样解释道："他们向我喝彩是因为他们都能看懂我的电影，而他们向你喝彩是因为没人能看懂你的理论。"

爱因斯坦的广义相对论变成了晦涩难懂的代名词：对智力的终极挑战。当然，为了让运动和引力的新定律满足哥白尼原理的要求，爱因斯坦所使用的数学语言一开始也难倒了他自己。他坦言他的数学不够好。他的天赋在于理解物理，而非数学技巧。但是如果遇到自己不会做的数学题，他知道谁会做。他的老同学、老朋友马赛尔·格罗斯曼（Marcel Grossmann，1878～1936）是一个天才数学家，即使现代数学最抽象的分支也难不倒他。格罗斯曼也认识到爱因斯坦拥有非凡的天赋，能够洞察自然规律的核心，他很愿意竭尽所能去帮助爱因斯坦。1912 年，在众多更有名的大学提供的教职面前，爱因斯坦接受了苏黎世联邦工业大学的教授职位，这样他就可以与在那里担任理论数学教授的格罗斯曼亲密合作了。

为了让爱因斯坦能够表达引力如何改变宇宙形状的想法，格罗斯曼为他介绍了一种新的数学工具。爱因斯坦渴求一种能“平等地”写下自然法则的方法，于是对任何人来说，自然法则看起来都一样。格罗斯曼为他展示了这一方法，这需要用到数学语言的一个深奥分支，叫做张量演算，这种数学能满足爱因斯坦所寻求的普适性。格罗斯曼又为爱因斯坦介绍了发生在复杂曲面几何中的重大进展，这正是史瓦西几年前开始探索的方向。不过，为什么爱因斯坦需要了解这些古怪的几何学呢？

牛顿所理解的空间是一个固定的巨大舞台，行星、彗星等天体的运动都发生在这个舞台上。天体可以你来我往，但无论舞台上占据着什么样的物质，发生了什么样的运动，空间都固定不变，而且不可改变。在爱因斯坦的理论中，空间的可塑性则要强得多。就像一大片橡胶板，随着上面物质和运动的不同，橡胶板的形状也在变化。在有大量物质聚集的地方，空间的曲率很大。而离物质越远的地方，空间也就越平坦，越没有形变。当一个物体从曲面上的一点向另一点运动时，会遵循其中的最短路径——也就是曲面上的“直线”。大质量附近的几何会发生严重的凹陷，物体遵循最短路径经过时会被拖向凹陷中心。这样等它经过之后，它的运动方向已经随着空间形状而发生了变化。用这种方法，爱因斯坦可以简单地认为引力来源于空间的形状。确实，根本不需要任何力，只需要有弯曲的空间。

乍听起来，这不过是在用另一种方式说明存在一种力。但事实并非这么简单。在牛顿的空间固定的旧世界中，你可以设想空间的固定舞台上有一个旋转的球。如果你站在舞台上离球不远处，它不会对你产生任何影响。爱因斯坦的世界就大不相同了。如果空间像一个形状可变的橡胶板，那么旋转的球就会扭曲它周围的空间①，你站在旁边时就会受到某种沿着旋转方向的拖曳作用。[4]这个区别可不小。

① 地球旋转造成的微弱时空拖曳效应于 2011 年被美国国家航空航天局（NASA）的引力探测器B证实。——译者注

为了把这种思想变成一种新的引力理论，爱因斯坦所做的工作就是寻找某些特殊的方程组——这种新的引力定律对于任何的物质和能量组合（不论是静止的还是运动的），都能告诉我们其空间有何种形状、时间以多大的速率流逝。而且这个定律还要告诉我们，这些组合应当以何种方式变化才能保证能量守恒。美国物理学家约翰·惠勒（John Wheeler， 1911～2008）把爱因斯坦的理论归纳成两句话："物质告诉空间如何弯曲。空间告诉物质如何运动。"

爱因斯坦用新的数学语言，也就是格罗斯曼教给他的张量分析把这些方程组写了出来。这确保了，不论观测者如何运动，旋转、加速、跳上跳下或是转圈圈，方程组的形式在所有的观测者看来都一样。不论他们的实验室处在什么样的运动状态，所有的观测者都会推导出相同的引力定律来。[5]

爱因斯坦的方程组有着优美的性质。也就是说，一个决定曲面几何如何变化的纯数学定理，等价于一组在所有自然过程中能量和动量都守恒的物理定律。值得我们注意的是，如果考虑质量很小、运动速度又远小于光速的情况，空间的几何形变也会非常微弱——这时爱因斯坦方程组就转化成了从前的牛顿引力定律。

一点题外话

谬误也有好处，而且越多越有好处。犯过错的人会进步，他们更值得信赖。为什么？他们不再有威胁。他们不会太较真。没犯过错的人总有一天会跌下悬崖，这是坏事，因为任何跌落的人都是累赘，他们落地时可能会砸在你身上。

——詹姆斯·丘奇（当代侦探作家）[6]

2000年，英国有一场大型比赛。人们通过投票，要选出过去一千年中最伟大的英国人。知名节目主持人和记者尽最大努力说服民众，把莎士比亚排在戴安娜王妃前面，把贝克汉姆排在达尔文后面。一家严肃的报纸想到了牛顿。这本来很有说服力——彩虹①、运动、万有引力、微积分以及其他累累成就——但有位记者没有被说服。"爱因斯坦证明牛顿的一些理论完全'错'了。"他如此写道。

① 牛顿用三棱镜做色散实验，证明白光不是单色光，而是由多种颜色的光混合而成。——译者注

这句评论透露出有关现代科学发展进程的一个误解。当爱因斯坦的引力理论出现，并被越来越精确的实验验证时，并不意味着牛顿的理论被扔进了垃圾堆。将来，爱因斯坦的理论也会被取代，但我们同样不会把它扔在一旁。

爱因斯坦的思想拓展了牛顿理论，于是我们就能了解引力场很强和运动速度接近光速的时候会发生什么。牛顿理论不能描述这种极端情况。但如果考虑爱因斯坦理论的极限情形，即物体运动速度远低于光速，并且引力场很弱的情况，那么爱因斯坦理论会变得越来越接近牛顿理论。也就是说，牛顿理论是爱因斯坦理论的极限近似。从应用范围和适用条件来讲，爱因斯坦理论可以取代牛顿理论；但作为对低速、弱引力场的日常世界的有效描述，牛顿理论还无法被取而代之。这就是为什么工程类的学生都要学习牛顿定律的原因，而我也敢预测他们在接下来的一千年内还是会学大致同样的内容。

关于沃纳·海森堡和保罗·狄拉克的量子力学取代牛顿力学，我们也要以同样的眼光看待。量子力学把我们所了解的力学适用范围拓展到了非常微小的空间和时间尺度下，包括质量极小的粒子和光子的运动。然而，当我们渐渐扩大事物的尺度时[7]，量子力学就会越来越像牛顿力学。

在物理学理论的演化中，新的理论会包含旧理论的成功之处和前人们的见解。它们拓展了理论的适用范围以及它们可以准确描述的情景。新理论没有推翻旧理论，也没有把它们扫进历史的垃圾箱。很久以前也许有那么几次，那是因为当时的理论几乎没什么证据支持，说服力微乎其微。如今，一个新理论必须要能解释所有已被解释的现象和一些还未被解释的现象，而且还要做出一些从没有人想到过的预言。

爱因斯坦静而实的宇宙

人们说歌剧已不是过去的歌剧了，他们弄错了。歌剧还是过去的歌剧。这才是出错的地方。

——诺尔·考沃德（1899～1973，英国剧作家、演员）[8]

1915年11月，爱因斯坦向科学界公布了他的新引力理论，论文发表在备受推崇的《普鲁士皇家科学院院刊》上。爱因斯坦花了十多年的时间才为引力的问题找到了满意的回答，并推导出了一组方程，可以描述质量和能量如何弯曲空间以及如何在弯曲的空间中运动。他的理论精确地预言了水星轨道近日点每世纪43角秒的神秘进动。自从1859年被法国天文学家奥本·勒维耶（Urbain Le Verrier，1811～1877）发现后，水星进动问题就一直困扰着天文学家。牛顿理论并不能直接解释这个现象。[9]在爱因斯坦看来，这是物理学和数学公式的有机结合——物理定律几乎是毫无预兆地从抽象数学中冒了出来，这使他在给一位朋友的信中兴奋地写道："几乎没有一个理解这个理论的人会不被其中的魔力所吸引。"[10]

一年半以后，1917年2月8日，正值第一次世界大战期间，爱因斯坦公开发表了第一次把他的新理论用在了整个宇宙上的成果。方程组的每一个解都表示一个可能的宇宙。然而似乎只有一个宇宙，那么我们该如

何剔除过多的宇宙呢？在这个问题上，爱因斯坦苦苦挣扎了很久。如果认为宇宙是无限的，他又无法确保方程组在无限远的地方仍然起作用。而如果认为宇宙有限，那么他就必须避免遇到空间的“边界”。

借助史瓦西曾关注过的数学进展，爱因斯坦发现了正曲率空间的重要性。它有限，但又像一个球的表面那样，没有边界。他还相信对称性：平均而言，宇宙必须在任何方向上、任何位置上都一样。所以尽管宇宙每一处的空间曲率都略有涨落，但这就像平静的海面，任何时候从大尺度看海面的任何位置、任何方向都一样。空间曲率所带来的一个有趣结果是，尽管各个方向看上去都一样，但这并不意味着你在宇宙的中心。如果你像蚂蚁一样在一个球面上爬，无论在哪里，你会觉得任何方向看上去都一样，而且球面并没有中心。[11]

但此后爱因斯坦并没有做出进一步的历史性飞跃，而是有点步履蹒跚。在做出所有这些简单的理论假设后，他却得不到任何静止的宇宙模型。所有可能的世界都必须随时间变化，无论在宇宙的什么地方，都会随着宇宙一起膨胀或收缩。空间可以弯曲，但是对1917年的爱因斯坦来说，空间作为恒星运动的场所必须是静止的、固定的。在他看来，能得出静止宇宙的唯一方法就是往他的方程组中引入一个可能存在的项，这一项的重要性此前被他低估了。

牛顿的引力理论告诉我们，两个物体之间的万有引力使它们相对彼此做加速运动。为了阻止这个加速，需要存在一种排斥的作用，于是物体受到的合加速度是

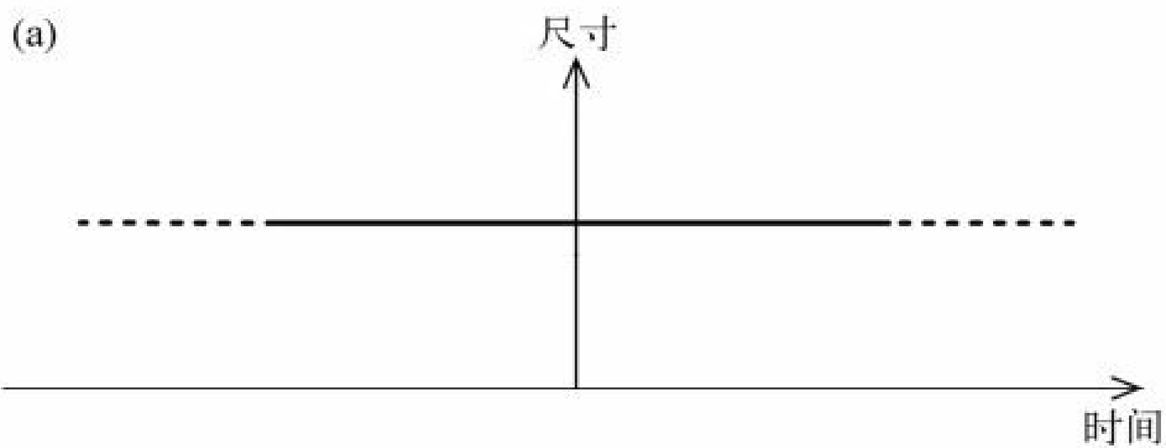
$$加速度=-万有引力+排斥力$$

爱因斯坦的理论允许排斥力存在，但并不要求排斥力存在：这是种额外的选项。这种排斥力看起来就像大自然还没利用的一个附件，因为我们在地球和太阳系中研究引力时并没有发现相关证据。物体离得越近，它的效果就越弱。但物体距离很远时，它的效果就增强了。[12]这说明在宇宙中的某一个距离上，排斥力和万有引力相等。这样大小的宇宙既不会膨胀，也不会收缩。这就是爱因斯坦的静态宇宙。

这种宇宙的空间曲率为正，所以是有限而无界的。如果我们画一个图，向上表示时间的流逝，空间垂直于时间的箭头（每一个时刻都是一个二维空间切片），那么对爱因斯坦宇宙来说，某些运动物体在时空中走过的路径正是这个圆柱体表面上一条蜿蜒的螺线（图 3.3）。如果向一个远离我们的宇宙飞船望去，它先是变小，然后又返回来变大了。光绕整个宇宙跑一圈所花的时间[13]取决于宇宙中物质的平均密度。[14]如果一个宇宙中物质的平均密度等于现在的空气密度，这个时间就是两天半。在这样的宇宙中，每当过去发出的光绕过整个宇宙又回到我们身边

时，我们就能看见两天半、五天、七天半、十天以前做过的事。

爱因斯坦的宇宙继承了史瓦西的静态空间的长处。他引入了一个惊人的宇宙模型——一个有限无界的、自无穷过去到无穷将来都存在的弯曲空间。这是他非凡方程组的第一个产物，但方程组想告诉他的信息却把他弄得很沮丧：宇宙并不想静止。后来，爱因斯坦回应说，“这是我一生中最大的错误”。


(a)
尺寸
时间

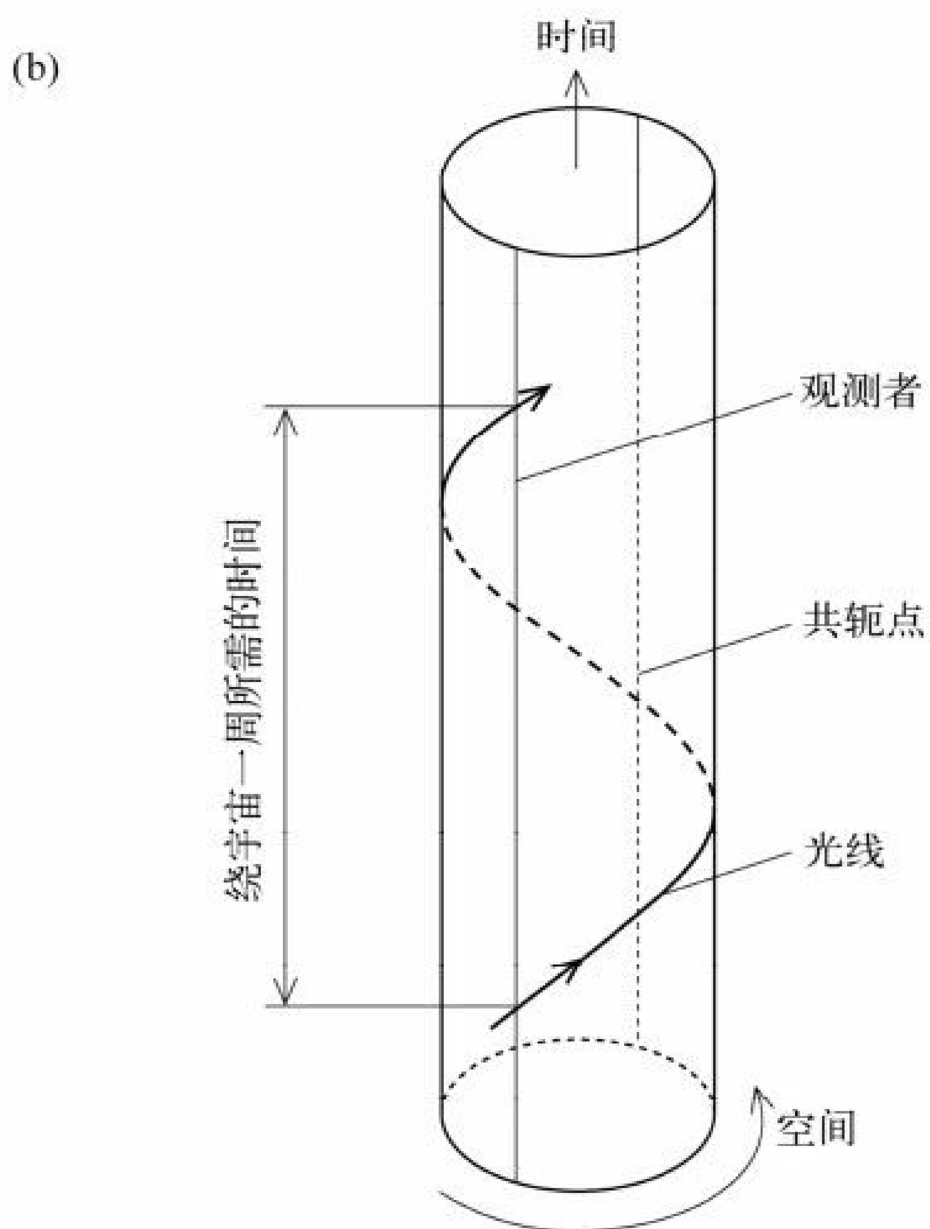

(b)
时间
观测者
共轭点
光线
绕宇宙一周所需的时间
空间

图3.3 (a)爱因斯坦静态宇宙模型的时空结构图。宇宙中自由运动的粒子之间的距离不会发生变化。(b)光在爱因斯坦静态宇宙中随着时间流逝的运动轨迹。这个轨迹是圆柱表面的一条螺旋线

德希特动而空的宇宙

我对宇宙非常感兴趣——我专门研究宇宙及其周边事物。

——彼得·库克（Peter Cook，1937～1995，英国讽刺作家）

紧随其后研究爱因斯坦方程组的是著名的荷兰天文学家威廉·德希特（Willem de Sitter，1872～1934）。得益于荷兰的战时中立，德希特见到了爱因斯坦，并和他一直保持通信。[15]他也同英国著名天文学家阿瑟·爱丁顿通信讨论科学问题，当时爱丁顿正负责组织1917年度在皮卡迪大街皇家天文学会举办的月度研讨会。正是在学会的第三场报告中，德希特向人们展示了爱因斯坦方程组的一个新的解。[16]

德希特保留了爱因斯坦的排斥力，但又设定宇宙的物质密度是零。当然，真实的宇宙不是空的。德希特假设物质密度非常低，因此产生的引力与爱因斯坦的排斥力相比完全可以忽略，这种排斥力用希腊字母Λ[①]表示。不像爱因斯坦的宇宙，德希特宇宙的空间几何是欧几里得的，因而是无限的。

① 读作拉姆达，英文拼写是lambda。——译者注

尽管很容易得到一个德希特宇宙，解释起来却没这么简单。遥远物体所发出的光的波长好像会被拉长，因此颜色会变红，而且光源距离越远，拉得就越长。这被称作“德希特效应”。1912 年，美国天文学家维斯托·斯里弗（Vesto Slipher，1875～1969）发现，一团遥远星云（现在我们管它叫“星系”）发出的光的特征波长存在明显的位移。[②]五年之后，他又报告了从其他二十多个星云的光谱中发现的红移现象。他不知道如何解释这些位移。现在德希特证明，他给出的爱因斯坦方程组的解能引起这种效应。进一步的研究给出了原因：德希特宇宙在膨胀。如果你标记了德希特宇宙中的两个点，它们就会加速分离，距离呈指数增长。图3.4画出了德希特宇宙中两个参考点的距离随时间变化的关系：随着时间流逝，距离在加速增长。

② 星光被分解成像彩虹一样的光谱，不同位置上的颜色，代表不同的波长。光谱的特征（亮线、暗线）发生移动，表明特征光波的波长发生了变化。——译者注

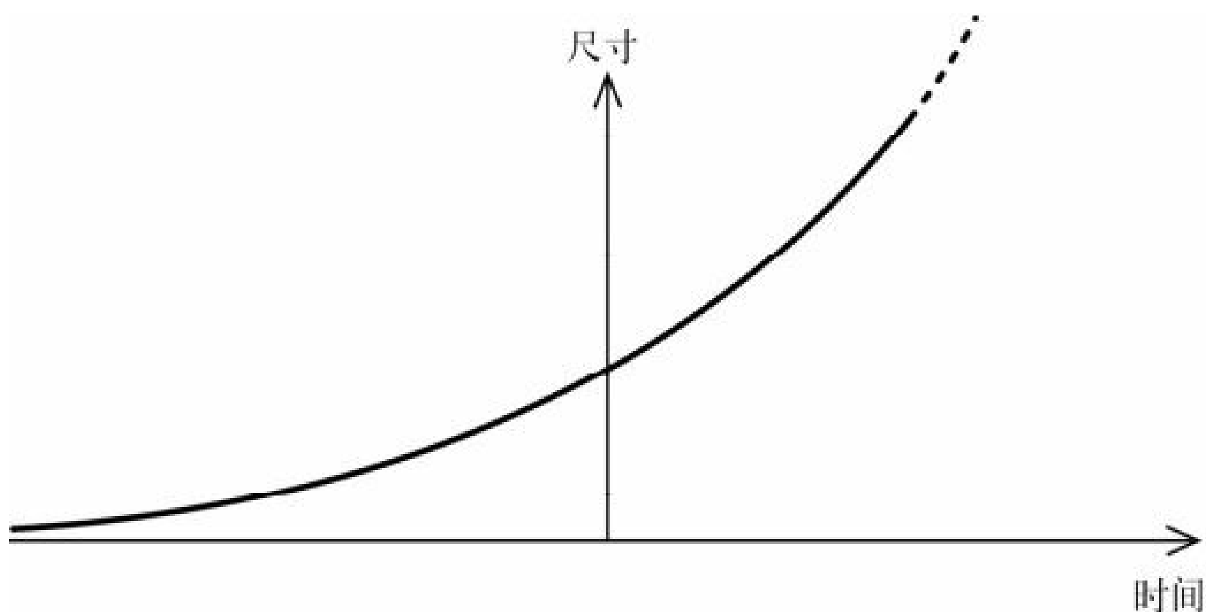


图 3.4 德希特的加速宇宙。自由粒子之间的距离会随着时间的流逝而呈指数增长

在这个膨胀宇宙的图景中，“德希特效应”有个很简单的解释。当一个退行的恒星发出某种光时，波长就会被“拉长”，我们接收到时就发现光的频率比发出时变小了。这个规律对所有的波都适用，特别是声波和光波。当光源靠近我们时，事情刚好相反。一个退行的光源变红，而靠近的光源变蓝。一个退行的声波音调变低，一个靠近的声波音调变高。这个现象叫“多普勒效应”，这是由奥地利物理学家克里斯蒂安·多普勒（Christian Doppler，1803～1853）在 1842年试图解释运动的恒星为什么有不同颜色时发现的。[17]我们更熟悉声音的例子。回忆一下凌晨三点街上的摩托少年从你卧室旁呼啸而过。那噪音很有特色：咦——哟。起先他朝你开过来，马达的音调升高（咦），然后他绝尘而去，声音从远去的马达跑回来，频率变小音调降低（哟）。

斯里弗观测到的现象正好与有些恒星从银河系一头急速远离我们，有些又从另一头靠近我们时所发出光波的情形相符。当然，斯里弗也许看到的只是一些飘移的天体，从某个方向飞来，路过我们之后又离开了。渐渐地，对银河系两头的天体的观测表明，这些天体都在远离我们，但斯里弗坚持用他的漂移假说来解释红移。①在当时没有理由能让他（以及任何其他人）相信宇宙在膨胀，或甚至只是考虑一下这可能意味着什么。

① 斯里弗在1917年发表的一篇题为“星云”［‘Nebulae’,Proc. Amer.

Phil. Soc. 56, 403(1917)］的论文中提到，他测量了分布在天空不同位置的 25 个星云的光谱，有一些在退行，主要分布在天空的一侧；有一些在靠近，主要分布在另一侧。这让他想起银河系内猎户座腰带上的一些恒星在退行，而天空另一侧的一些恒星在靠近。斯里弗猜测这可能是因为我们所处的星系在宇宙中漂移。现在我们知道这个结论只是巧合，因为斯里弗分析的样本太少了。——译者注

在德希特宇宙的数学模型中，宇宙的全部空间确实都在爱因斯坦拉姆达力的作用下加速膨胀。可惜没有人想到要把这件事与斯里弗在1917年的观测联系起来，连德希特都没想到。直到1921年，德希特才知道斯里弗观测的螺旋星系中，有22个在远离我们。没人知道它们离我们有多远，而且仍有可能是局部的随机运动引起了这种现象，而不是整个宇宙的膨胀。德希特不愿意从斯里弗的观测中做出任何肯定的结论。他第一个得到了膨胀宇宙模型，但宇宙的膨胀却犹抱琵琶半遮面，还得一阵子才能露出它的庐山真面目。

德希特的宇宙对我们现在的宇宙学研究非常重要。请注意它的一些特点。德希特宇宙的体积会变得越来越大，没有开始，也没有结束。相反，当你逆着时间往回追溯时，它会变得越来越小，却永远不会变成零，也不会有大小为零而物质密度为无穷大的明显开端。它的膨胀速率恒定不变。如果在某一时刻掉进这个宇宙里，你根本无法确定你的具体时刻，因为未来和过去没什么区别。你能观察到的所有事物都永远一个样子。在德希特的世界中，历史并不是一个重要主题。

弗里德曼动而实的宇宙

我们有爱因斯坦空间、德希特空间、膨胀的宇宙、收缩的宇宙和振荡的宇宙。事实上，纯粹的数学家仅仅通过写下方程就可以创造出宇宙。如果他信奉个人主义，他甚至可以拥有一个他自己的宇宙。

——约瑟夫·约翰·汤姆逊（1856～1940，英国物理学家）

爱因斯坦是个物理学家，德希特是个天文学家。在这场发现新宇宙的竞赛中，紧接着出场的新选手虽然后来鼎鼎有名，但在当时却只是个名不见经传的年轻数学家、气象学家，他就是来自圣彼得堡的亚历山大·弗里德曼（Alexander Friedmann，1888～1925）。[18]他很幸运，当他还是个物理系的年轻学生时，就参加过奥利地杰出物理学家保罗·艾伦费斯特（Paul Ehrenfest，1880～1933）的量子理论和相对论课程，艾伦费斯特在去莱顿大学之前于1907～1912年间在圣彼得堡大学任教。弗里德曼毕业后一直和艾伦费斯特保持联系。他先是在巴夫洛斯克观测站当一名气象学家，后来又去莱比锡成为现代气象学理论奠基人威廉·比约克内斯（Vilhelm Bjerknes，1862～1951）的研究生。第一次世界大战期间，他在奥匈前线从事弹道学工作。[19]战争结束后，他重新拾起学术，

并立刻有了进展。他的研究涵盖数学、矿物学和大气科学。最后，弗里德曼在1918年成为彼尔姆国立大学的数学和物理学教授，这个大学是圣彼得堡大学的新分校（outpost）。在那儿，他遭受了内战的折磨，彼尔姆市先后被反苏的“白军”和托洛茨基的红军占领，这使得他的绝大部分同事相继离去。1920年，弗里德曼去了圣彼得堡的地球物理学观测站，并在那里开始学习爱因斯坦新的广义相对论。弗里德曼的研究领域极为广博，既有纯数学理论研究，也有引人注目的高空气球飞行，这是为了研究人体的高空反应。在1925年的一段时间里，他同一位同事一起曾一度保持着高空气球飞行的世界纪录，到达7400米的高度。可惜几个月后他就去世了，年仅37岁，死因似乎是斑疹伤寒。[20]

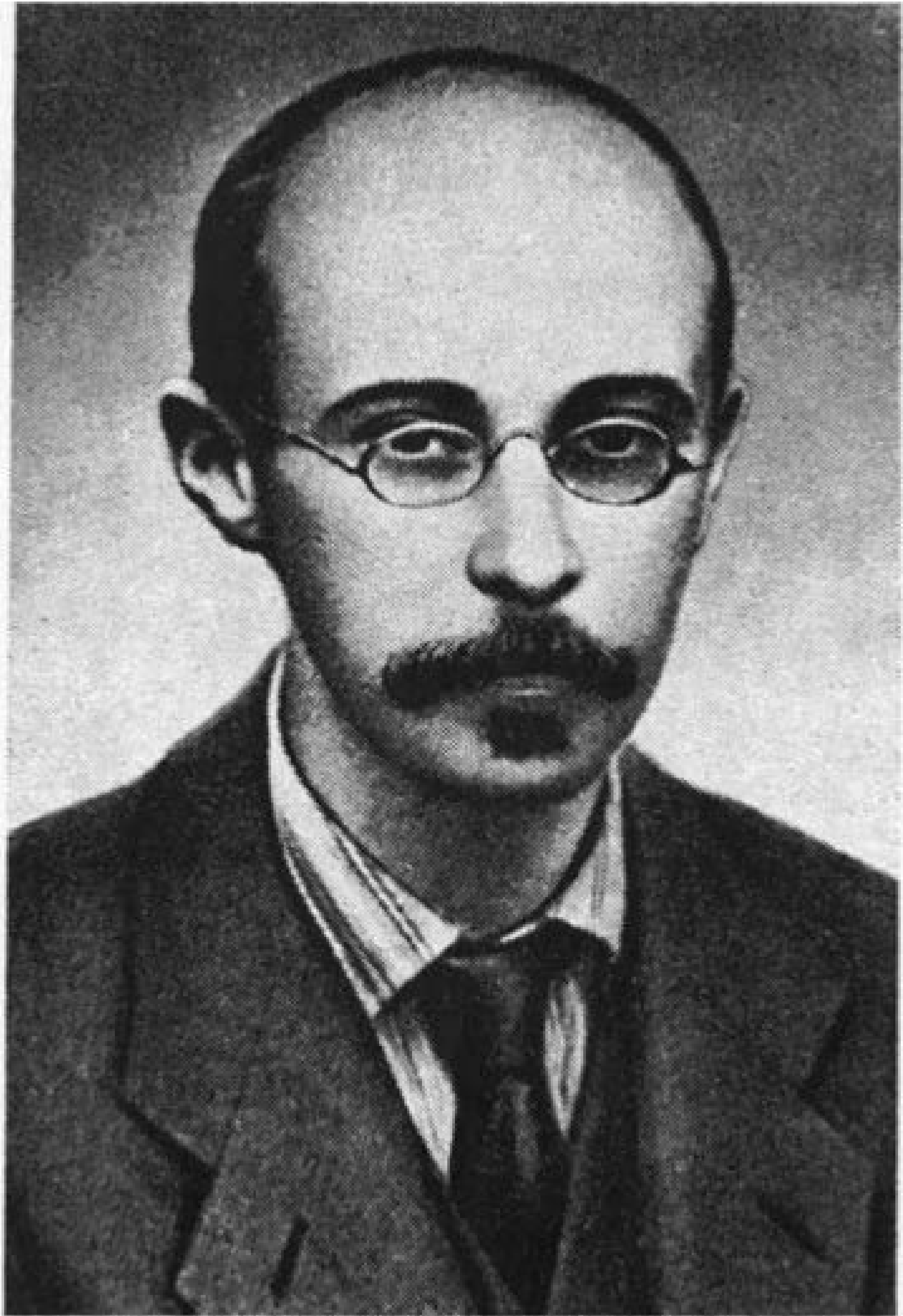


图3.5 亚历山大·弗里德曼

弗里德曼已经非常详细地学习了爱因斯坦方程组背后的艰涩数学，并着手寻找比爱因斯坦和德希特更一般的解，同时保留他们关于宇宙各处各方向都一样的假设。从弗里德曼1922年和1924年的两篇文章以及1923年撰写的《在时空中的世界》一书中看来，他应该听说过德希特和爱因斯坦的宇宙，但好像不知道斯里弗发现遥远恒星的光波红移现象。他研究爱因斯坦方程组的方法就是以数学家的眼光寻找新的解，而且他找到了。

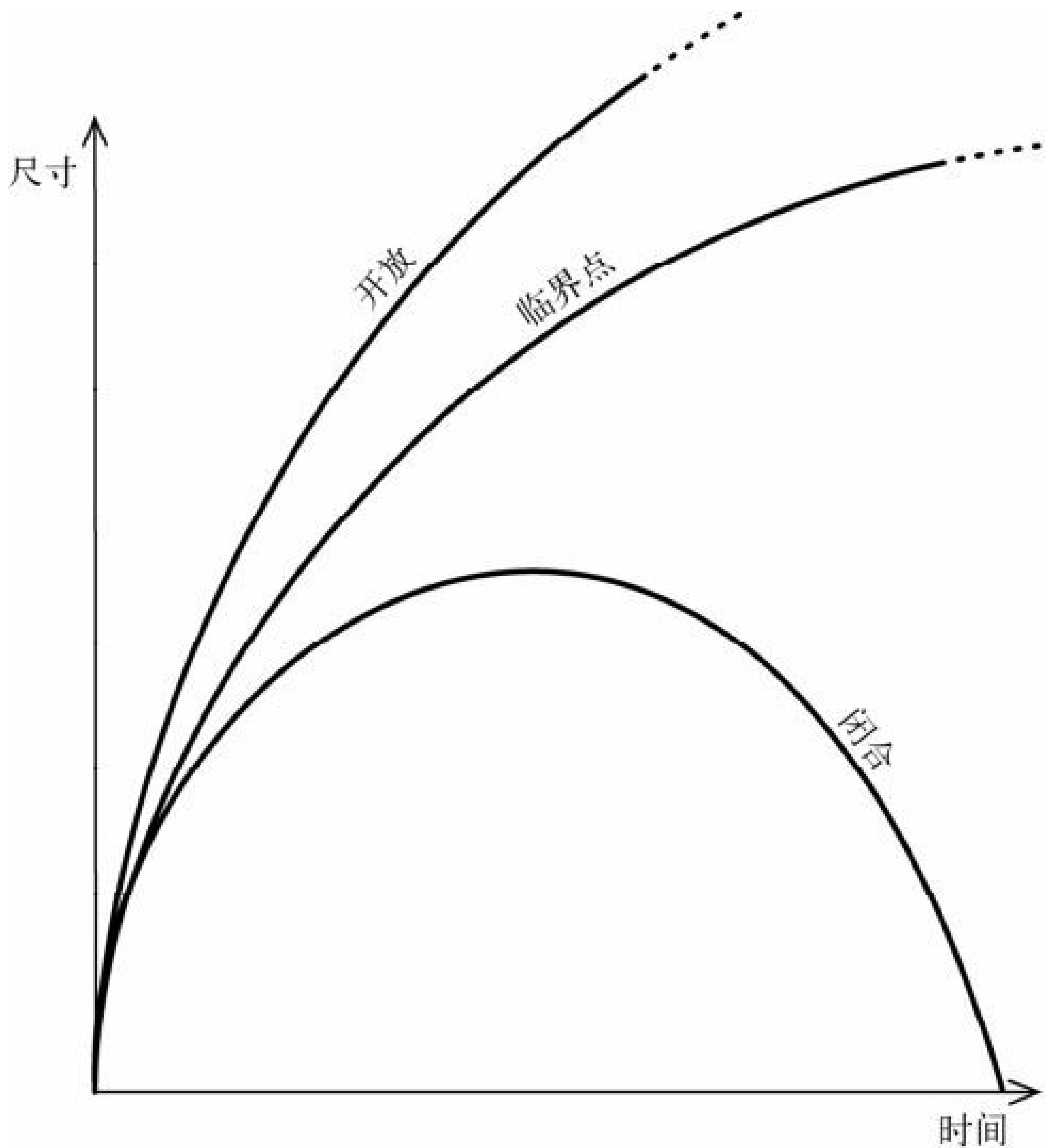


图3.6 弗里德曼的膨胀和收缩的宇宙

首先，他发现了一种有限的“闭合”宇宙。这种宇宙有正曲率的空间，从有限的过去开始膨胀，到一个极大值后又开始收缩，直到有限未来的终点（图3.6）。这是个膨胀的宇宙，包含普通的无压强物质①。宇宙中物质的质量是有限的，宇宙的体积和寿命也是有限的——弗里德曼甚至估计，如果这种宇宙循环一周的时间跨度大约是 100亿年的话，它的质量大概是太阳质量的5×10^{21}倍。[21]这种宇宙始于一种密度无穷大、后

来被称作大爆炸的过程，又会在与之相反的极端大塌缩过程中收缩到一个终点。弗里德曼在他的书中提出，如果在时间上向前（和向后）延拓这个解，这种宇宙就会周期性地膨胀和收缩，在一个无穷的序列中振荡（图3.7）。他这样写道：

① 这里说的压强指相对论性压强，除了光，生活中遇到的大部分物质的压强都可以当作零。——译者注

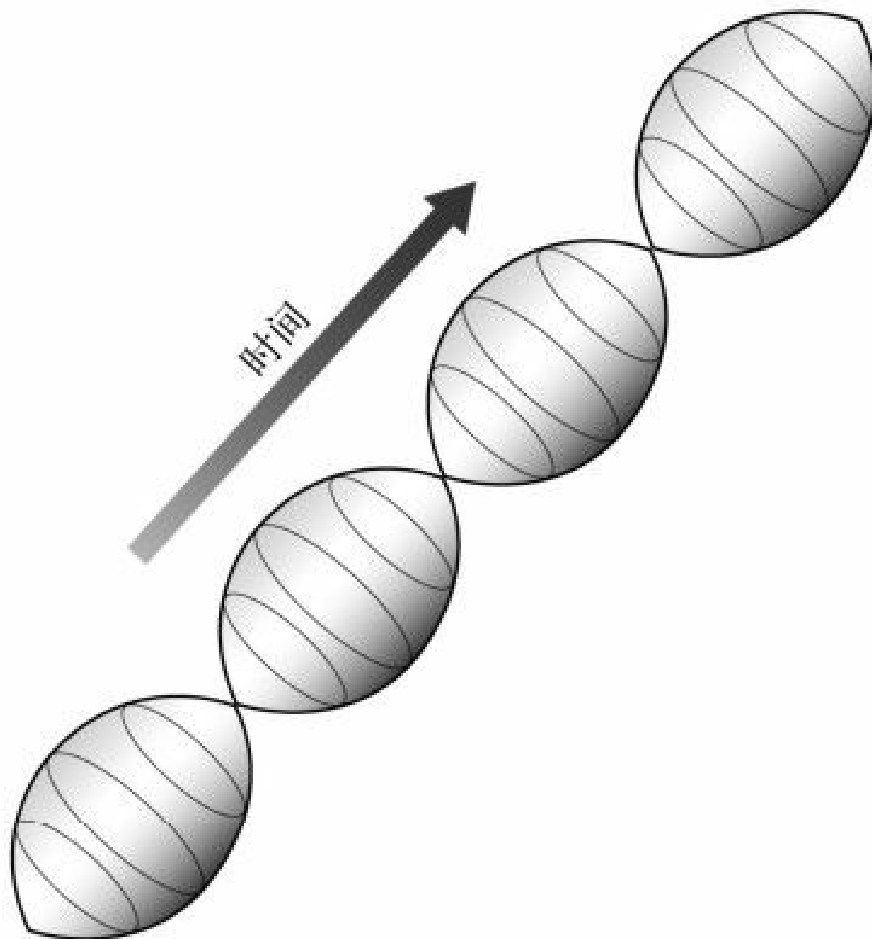


图3.7 弗里德曼的振荡宇宙

可能存在这样的情况，世界的曲率半径……周期性地变化：宇宙收缩到一点（变成虚无），然后半径又从一点变大，到某个值以后，曲率半径又缩小，变成一个点，依次类推。这让人情不自禁地想到印度神话中的生命轮回。同时也有这样一种可能性，世界从虚无中创生。但这一切目前只能算作天方夜谭，因为太缺乏天文学观测，无法进行检验。[22]

紧接着，弗里德曼又发现了爱因斯坦方程组的另一类解，有着“开放”的空间、负的曲率，就像一个马鞍的形状，因此这种宇宙的体积是无限大的，会从无限的过去开始膨胀，直到无限的未来。[23]图 3.6 和图

3.7中也画出了这类解。

所以，是弗里德曼第一个发现了爱因斯坦方程组允许膨胀或收缩的宇宙包含普通物质，如行星和恒星。他并不是受到天文学观测启发之后才产生这个想法，也没有拘泥于宇宙的开始或结束所包含的物理学意义。有趣的是，他描述了一种从虚无中产生（又在虚无中消失）的宇宙。但弗里德曼时运不济，人们完全没有注意到他的重大发现。他的论文发表在当时的顶级期刊《物理学杂志》上，爱因斯坦在上面也发表过一些成果，但完全没有提到他的新“宇宙”。更倒霉的是，爱因斯坦认为弗里德曼有一些关键性的计算错误，这些膨胀收缩的新宇宙根本不是他的方程组的解，为此还在期刊上发表了一篇短文来特别进行了说明。

幸运的是，1923年5月，弗里德曼在圣彼得堡的一位同事尤里·克鲁科夫（Yuri Krutkov）来到莱顿，在那儿见到了爱因斯坦，并且成功地说服了他，使他相信弗里德曼的计算是正确的：爱因斯坦的方程组真的有非静态的宇宙解。爱因斯坦迅速行动，在期刊上又发表了一篇短文说，在同弗里德曼通过信并与克鲁科夫讨论过之后，他意识到自己犯了一个计算的错误，而弗里德曼的解是“正确而清楚的”。这些解表明，场方程除了有静态解之外，还有随时间变化的解。[24]有意思的是，在投给期刊的那篇短文的手写草稿中，爱因斯坦在结尾加了一句总结陈词说，抛开这些解的正确性不说，“在其中几乎找不到物理学意义”。幸好在短文发表前，他把这句话删掉了。

弗里德曼的寿命不长，也没有通过天文学观测来解释他的发现。对他来说，这只不过是一些数学公式。他的一位同事弗拉基米尔·福克（Vladimir Fock）说，有一次弗里德曼告诉他：“他的任务就是暗示爱因斯坦方程组可能存在这样的解，并且物理学家们可以像他们所期待的那样研究这些解。”[25]但如今，没有一个名字比弗里德曼和宇宙的关系更紧密：如果在网上搜索一下“弗里德曼的宇宙”，你就会得到数百万个结果。

勒梅特的宇宙

世界的演化可以比作一场刚刚结束的烟花表演：几缕幽幽的红光、余迹和烟雾。站在一堆早已冷却的灰烬上，我们看到众多恒星逐渐黯淡，而我们正试图追忆那些逝去的辉煌。

——乔治·勒梅特[26]

1923年，乔治·勒梅特（Georges Lemaître，1894～1966）被圣龙博神学院（Maison Saint Rombaut）任命为天主教牧师，之前他还曾在“一战”中获得十字勋章。他本来是在比利时的天主教鲁汶大学注册学习工程学，但战争中断了他的学业，直到1920年，他才从数学系毕业。1923～1924 学年，他获得了去国外的奖学金，来到剑桥的圣埃德蒙学院[27]，成为大学天文台里阿瑟·爱丁顿的访问学生。当时，爱丁顿恐怕是世界上最有造诣的天体物理学家，拥有一系列卓越成就，诸如提出了恒星的

工作机制，发展了银河系内恒星的运动理论，以及领导了1919年那场著名的科学探险——去西非附近葡萄牙所属的小岛普林西比上，探测遥远恒星的光在经过太阳引力场时有没有弯曲，以检验爱因斯坦的预言。他也是第一个撰写英语论文介绍爱因斯坦的广义相对论的人。

爱丁顿拥有传奇般的数学能力，并且迅速而深刻地理解了爱因斯坦的引力理论。在“一战”期间，他作为皇家天文学会的秘书长，维持了同欧洲大陆科学家的科学交流。作为一名虔诚的贵格派教徒，爱丁顿没有加入那场战争。1917年，由于公开表示自己出于宗教信仰而拒绝服兵役，他险些被关进监狱颜面扫地，所幸有皇家天文学家弗兰克·戴森爵士（Frank Dyson）出面调停。利用和海军部的良好关系，戴森安排了一份关于推迟爱丁顿服役的协议，约定如果到时战争结束了，爱丁顿将领导海军部的项目里两支探险队的其中一支，于1919年5月29日观测一场日全食，以检验爱因斯坦的广义相对论。

勒梅特在剑桥交流的一年并没有向爱丁顿学习宇宙学，而是把精力投入到了深刻地学习广义相对论之中。离开剑桥后，他开始读博，师从哈佛天文台的著名天文学家哈罗·沙普利，并于1927年7月获得隔壁的麻省理工学院的学位（哈佛天文台到 1929年才开始授予博士学位）。爱丁顿对勒梅特的天才和数学能力留下了深刻的印象，并在信中向其他科学家强烈推荐他。勒梅特交游广泛，与人和善，在他的学术生涯中同每个人都能和谐相处。这无疑促进了合作与交流。

在波士顿的时候，勒梅特想通了红移的问题，并且已经阅读了爱因斯坦早期关于静态宇宙模型的论文。他和爱丁顿都没读过弗里德曼的文章。1927年，关于爱因斯坦理论预言的简单宇宙，勒梅特做出了最透彻的研究工作。他比爱因斯坦、德希特和弗里德曼走得更远，除了恒星和星系之外，又向宇宙中引入了有压强的辐射。同时他也试图用膨胀宇宙中的多普勒效应来解释斯里弗所看到的红移。

1927年，勒梅特完成了一篇重要的论文。它先是用法语写成，发表在比利时的不知名期刊上。这篇文章第一次把爱因斯坦方程组的膨胀宇宙解及其物理学解释，同遥远星光由于多普勒效应而产生红移的计算结合了起来。[28]就像他的其他研究工作那样，勒梅特思路非常清晰，没有多余的公式，并囊括了所有物理学的关键点。[29]他意识到，宇宙没有中心也没有边界，可以有限也可以无限，以及就能量守恒而言，爱因斯坦方程组有一个简单的物理学解释。他甚至从42个星系的红移和距离的观测数据中，算出了当前宇宙的膨胀速率，第一次给出了所谓“哈勃常数”H的数值（625千米每秒每百万秒差距[①]，与两年后哈勃的结果相近）。他计算了星系退行速度 v和距离 r的比值，并根据多普勒效应第一次推导出了哈勃定律（$v=Hr$）。

① 意思是说，每过1秒钟，宇宙中每相距326万光年的两个点之间的距离都会再增大625千米。——译者注

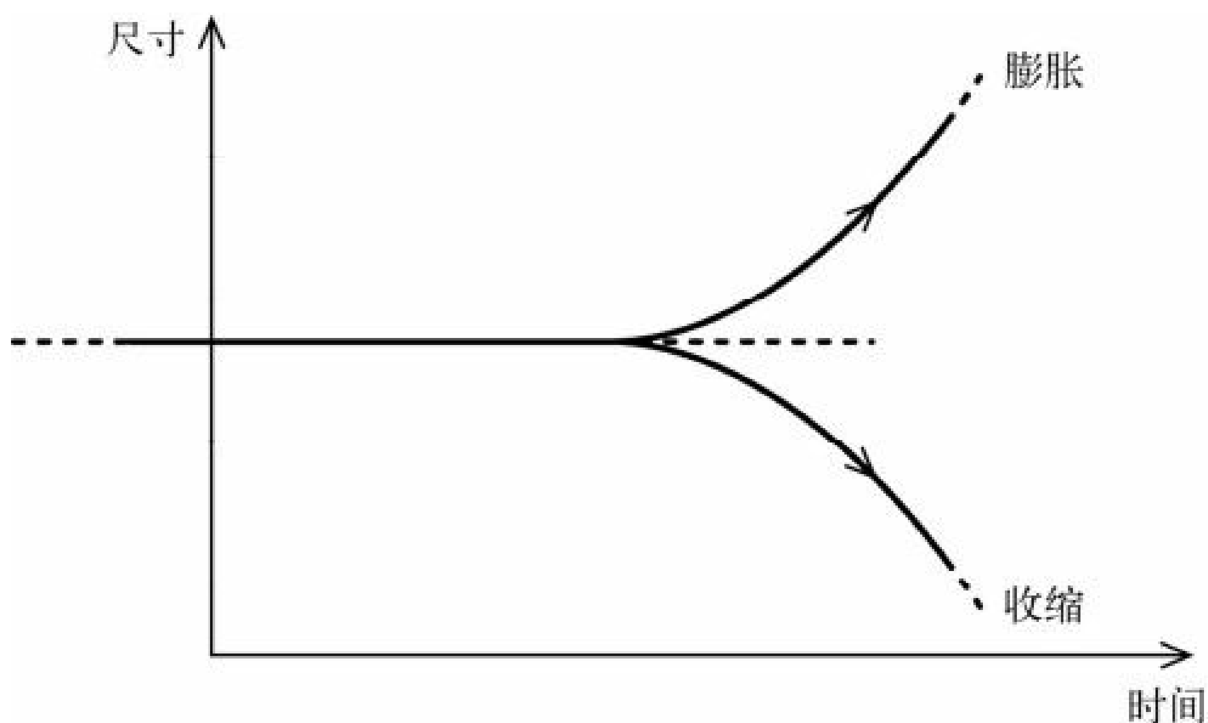


图3.8 静态宇宙的不稳定性。勒梅特和爱丁顿都证明，略微移动静态宇宙中的物质，就会导致宇宙开始膨胀或收缩

1929年，哈勃发表了自己的观测数据以验证哈勃定律。1930年，德希特紧随其后重新分析了这些数据（这让哈勃非常不爽，因为他把数据看作是自己的成果，发表后被德希特引用也不行）。哈勃从来没有真正赞同膨胀宇宙的全部物理学内涵，也没有专门用观测数据来验证理论模型。他把遥远星系的速度称作“视”速度，而把解释的工作留给了其他人。[31]无论如何，德希特和勒梅特都以极大的热情发展了膨胀宇宙模型。

勒梅特的工作明确了爱因斯坦静态宇宙和德希特指数膨胀的空洞世界的地位。勒梅特证明，爱因斯坦的世界不稳定。如果宇宙开始于一个静止的状态，那么其中的任何扰动或运动都会把它带入一个逐步升级的膨胀或收缩的状态（图3.8）。这等价于把宇宙立在一个针尖上。

这个发现虽然似乎没受到爱因斯坦的热烈欢迎，却成为勒梅特的宇宙名声大振的催化剂。1930年，他原来的导师阿瑟·爱丁顿证明了先前的怀疑，即爱因斯坦的静态宇宙在微小扰动下会不稳定。他向爱因斯坦的解中引入微小的密度不均匀，并发现所有不均匀性都会随时间长大。



图3.9 乔治·勒梅特和阿尔伯特·爱因斯坦（摄于1933年）

当时他没有注意到勒梅特的工作，但在文章发表后[32]，爱丁顿很快就收到了以前学生的来信，勒梅特指出爱因斯坦宇宙的不稳定性早在他1927年的文章中就已经被（用另一种方法）证明过了。爱丁顿震惊了，他已经忘了文章的这部分内容，也忘了勒梅特做过的这些计算的重要性。他迅速做出反应，于1930年7月给《自然》杂志写了一封信，提到勒梅特被忽略的工作，并安排将1927年那篇文章的英文版翻译发表在了1931年的《皇家天文学会月报》上[33]。结果，勒梅特成了当时最著名的宇宙学理论家。在1930年10月份的一次会议上，他才得知了弗里德曼先前所做的数学工作，但他还是在1931年的英文版翻译中增加了一条引用。


M

图3.10 勒梅特正在给学生上天体力学课

1957年，爱因斯坦去世两年后，勒梅特接受了一个关于他和爱因斯坦会面内容的采访。他披露了在1927年索尔维会议的私下场合中，爱因斯坦赞扬他的文章具有优美的数学形式，但从物理学的角度来讲，这些非静态的宇宙学“很讨厌”。[34]勒梅特仍然感觉到，爱因斯坦并没有真正领会退行星系的最新观测数据对他理论中的膨胀宇宙解具有什么样的意义。然而，1933年，在帕萨迪纳听了勒梅特关于此问题的报告后，爱因斯坦开始相信勒梅特方法的简洁性，并把他关于膨胀宇宙有一个炽热开端的图景称为对宇宙行为作出的最“美”的解释。

勒梅特的宇宙与爱丁顿在研究爱因斯坦宇宙不稳定性时发现的宇宙类似，因此它也常被称作爱丁顿–勒梅特宇宙（图3.11）。它开始于无限过去的一个静止状态，然后逐渐膨胀，直到在过去的某个有限时刻，膨胀的过程变得明显起来。[35]然后它继续膨胀，经过很长的一段时间后，越来越像德希特的指数膨胀宇宙。它有爱因斯坦的拉姆达排斥力；它的空间有限，曲率为正，并且永远膨胀。

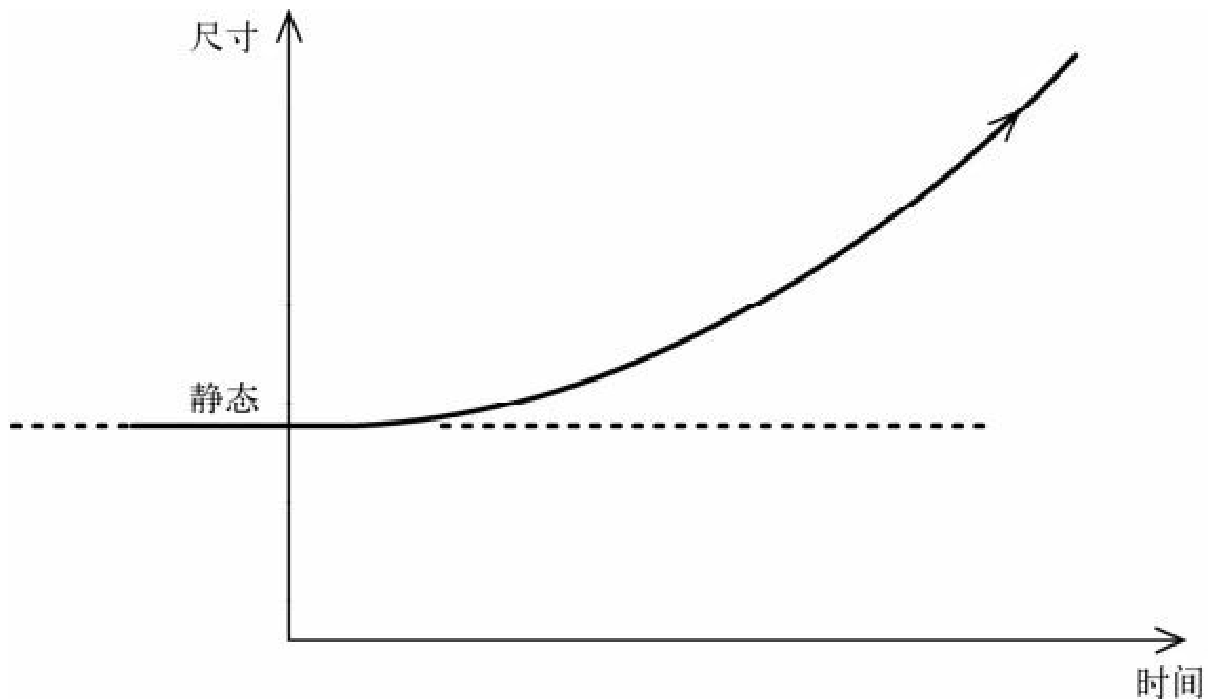


图3.11 爱丁顿–勒梅特宇宙

过去经过了无限长的时间意味着宇宙没有开端。这没有引起爱丁顿

和勒梅特的忧虑。事实上，爱丁顿把这种情况看作是大自然的一种特点，给了宇宙“无限时间用于开始”，因为“没有什么事情需要着急”。[36]他觉得一个突然开始的宇宙，就像勒梅特的宇宙，“令人不快”，只不过是个“烟花理论”，相比之下他的“沉稳”观念更能满足人们渴求安宁的心灵。爱丁顿说，尽管过去的时间有无限长，但有故事的过去是有限的，因为在无限远的过去，宇宙的状态过于接近完全的热平衡，无法产生大量的熵（图3.13）。因此，他相信这种类型的宇宙在其无限的历史中，不会因产生大量的熵和无序度而遭遇热寂。它的几何年龄很大，但热力学年龄还很小。

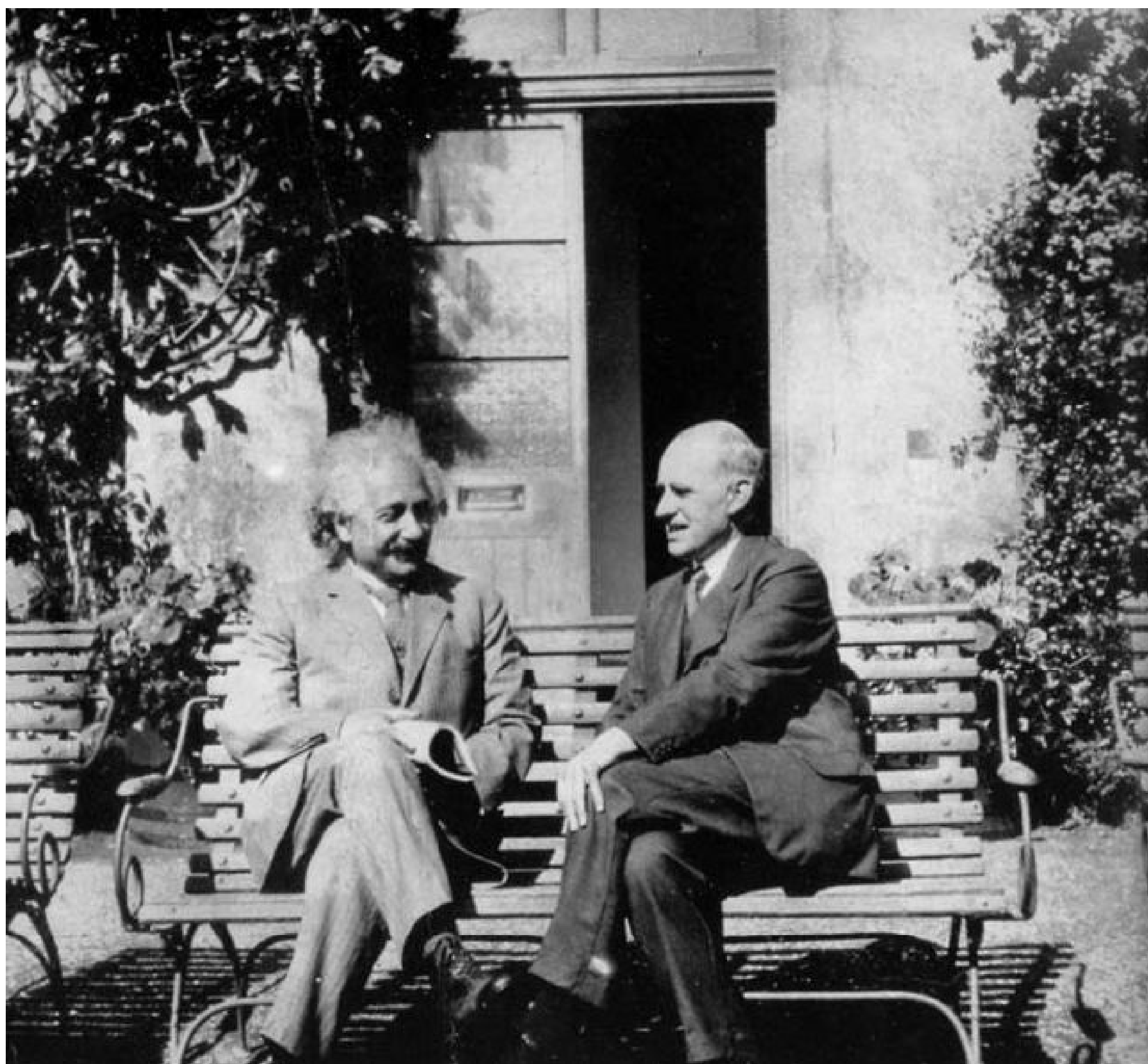


图3.12 1930年，阿瑟·爱丁顿和爱因斯坦在剑桥大学天文台的空地上聊天，当时爱丁顿和他姐姐住在那里

也许有人认为勒梅特被自己的宗教观念左右，所以才偏好有开端的宇宙。但实际上并不是这么回事。勒梅特把科学和宗教分得很清楚，将它们看作不相关也不矛盾的事情；两者是两套平行的体系，是对世界的不同解释。在他看来，《圣经》里没有科学，而从科学中寻找宗教教诲就如同从二项式定理中寻找天主教教规一样。[37]在他生命的后期，作为教皇科学院院长，勒梅特这样描述他的膨胀宇宙理论：

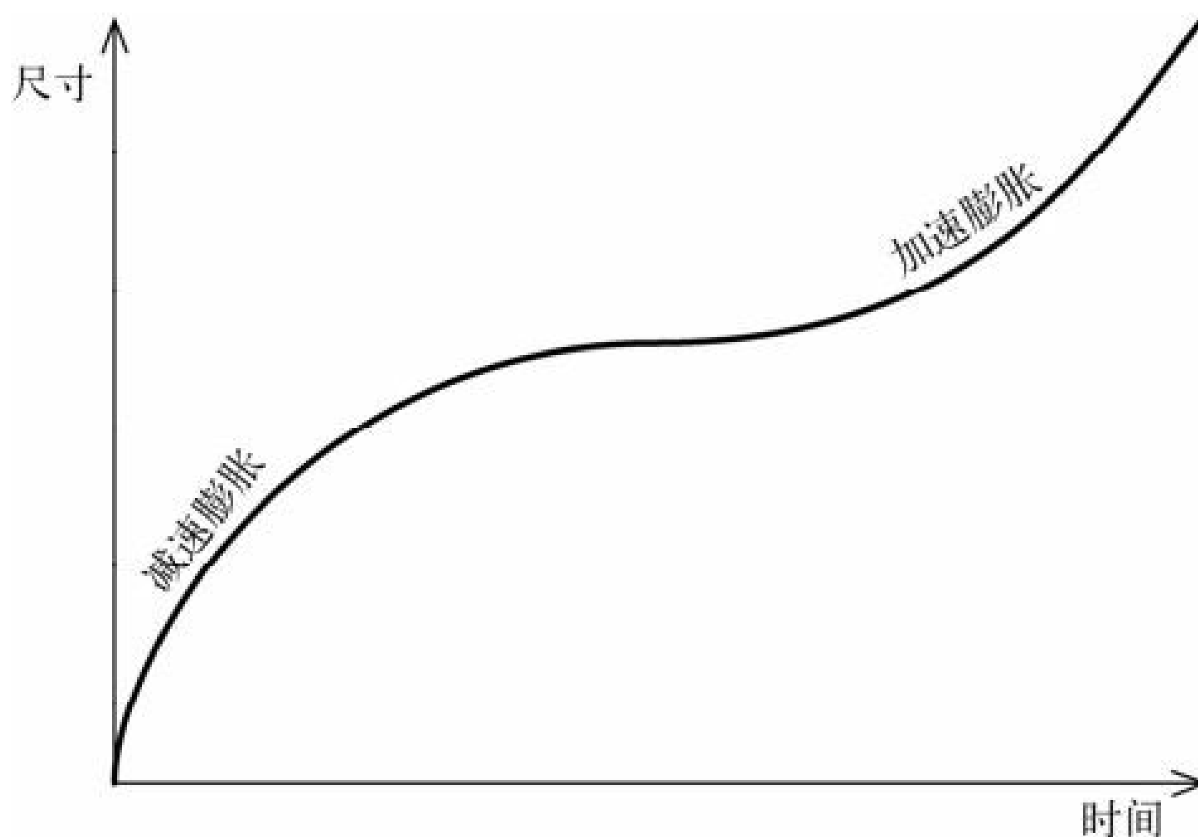


图 3.13 勒梅特宇宙的膨胀就像一个弗里德曼宇宙，始于一场大爆炸，有一段时期很像静态宇宙，然后开始加速膨胀，变得越来越像德希特宇宙

在我看来，这样的一个理论仍然完全与任何形而上学或宗教问题无关。它允许唯物主义者排斥任何超验的存在……对那些信徒们来说，它也要求他们对上帝避免任何“近之则不逊”……这与以赛亚所说的隐蔽的上帝相合，甚至在宇宙开端时也是隐蔽的。

尽管如此，他好像还是更喜欢他在1927年论文中发现的宇宙模型。那个模型确实有个有限的过去，并且从一个炽热的开端开始膨胀，先是减速膨胀，待宇宙学常数的排斥力占了主导，超过牛顿万有引力之后，逐渐变成加速膨胀，最后一直加速下去，走向德希特的指数膨胀宇宙（图3.13）。它有正的空间曲率、正的宇宙学常数，就像爱因斯坦的静

态宇宙，但其中的排斥力比爱因斯坦选择的特殊数值稍稍大了一点，因而宇宙永远在膨胀。

结果，勒梅特的宇宙是关于现实宇宙的最精确的描述，我们的宇宙有137亿岁，并于45亿年前从减速膨胀变为加速膨胀。

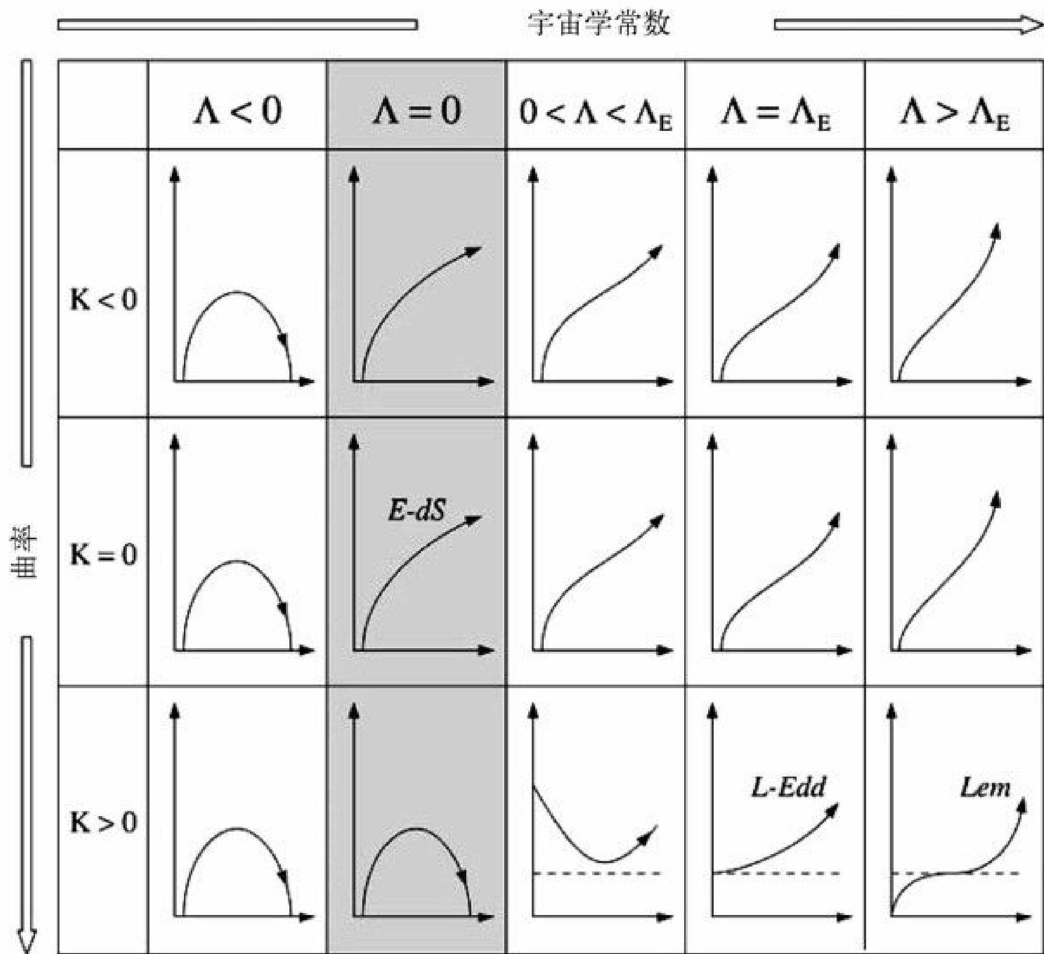


图 3.14 各种可能的宇宙模型，展示了满足空间曲率（K为正、负、零）和宇宙学常数各种可能取值的宇宙中，距离可能随时间的变化关系。当宇宙学常数取到某个特殊值，即 ${}_{E}$λṪλ 且空间曲率为正时，就是爱因斯坦的静态宇宙（虚线所示）。这种静态宇宙不稳定，最终会开始膨胀或者收缩，就像图中L-Edd模型或Lem模型一样

勒梅特的精辟分析让宇宙学家们得以直接研究各向同性、均匀膨胀宇宙的集锦。一共只有两个变量：空间的曲率可以为正、负或零（欧几里得的），而爱因斯坦引入的宇宙学常数产生的力可以为排斥（正）、吸引（负）或零。图3.14是各种可能得到的宇宙的集锦，1967年首次由

爱德华·哈里森（1919～2007）总结成一张表。[38]

爱因斯坦和德希特的宇宙

我自己当时并不觉得这篇论文有多重要，但德希特很热衷。

——阿尔伯特·爱因斯坦

你会读到爱因斯坦和我写的这篇论文。我自己并不觉得结论有多重要，但是爱因斯坦好像觉得很重要。

——威廉·德希特[39]

1932年的早春，爱因斯坦和德希特合作发表了一篇两页长的短文，试图简化宇宙学的研究。[40]勒梅特的宇宙集锦展示了许多可能的膨胀宇宙，其中一些永远膨胀，而另一些，例如弗里德曼最早发现的宇宙，最后会转而收缩。爱因斯坦和德希特当时都在帕萨迪纳的加州理工学院访问（图 3.15 是他们在一起工作），他们指出在大量均匀和各向同性的宇宙模型中，存在一个最简单的可能。

如果空间的曲率取作零（这样空间几何就是欧几里得的）、宇宙学常数设为零（爱因斯坦渴望加入他先前的发明）而物质的压强也设为零，结果就有了一个简单的宇宙。爱因斯坦–德希特宇宙从过去的某个有限时刻开始膨胀，直到永远（图3.16）。[41]

这篇文章只不过是在已有工作的基础上进行的简单推导，如果不是两大知名作者联合署名，怕是会由于原创性不够而不能发表。实际上，这个模型不稳定，就像爱因斯坦的静态宇宙那样。也就是说，如果曲率不是绝对为零，宇宙的膨胀就会逐渐脱离爱因斯坦–德希特的轨迹，越来越快，或越来越慢，直至开始收缩，如图 3.6。爱因斯坦–德希特宇宙是图中间那条线，而其余两个分别是闭合和开放的，宇宙学常数为零，随着时间变化而逐渐脱离中间的轨迹。爱因斯坦和德希特的想法的新奇之处在于，这个模型在随后的六十年中都是对现在宇宙膨胀行为的最佳描述，虽然他们俩没觉得有这么重要。①当时人们观测到的宇宙膨胀速率仍然没有偏离这个特殊取值，说明不稳定性还没来得及发展壮大。但是宇宙已经膨胀了一百三十多亿年了，这又说明宇宙开始膨胀时的状态必须极为接近爱因斯坦–德希特宇宙的特殊状态。这个奇怪的事情被称作“平坦性疑难”②，是阿兰·古斯（Alan Guth）在 1981年创建著名的暴胀宇宙理论的动机之一。我们将在后面的第9章说到他。

① 受到观测手段的限制，人们一度猜测宇宙学常数的观测值是零，因此作者说这个模型曾经是最佳描述。——译者注

② 据《热大爆炸宇宙学》，如果不考虑暴胀，那么自宇宙的普朗克时期开始，空间对平坦的偏离程度应该增大了58个量级。但今天观察到的空间仍然是非常平坦的。这反过来对宇宙的初始条件提出了极为苛刻的要求。——译者注

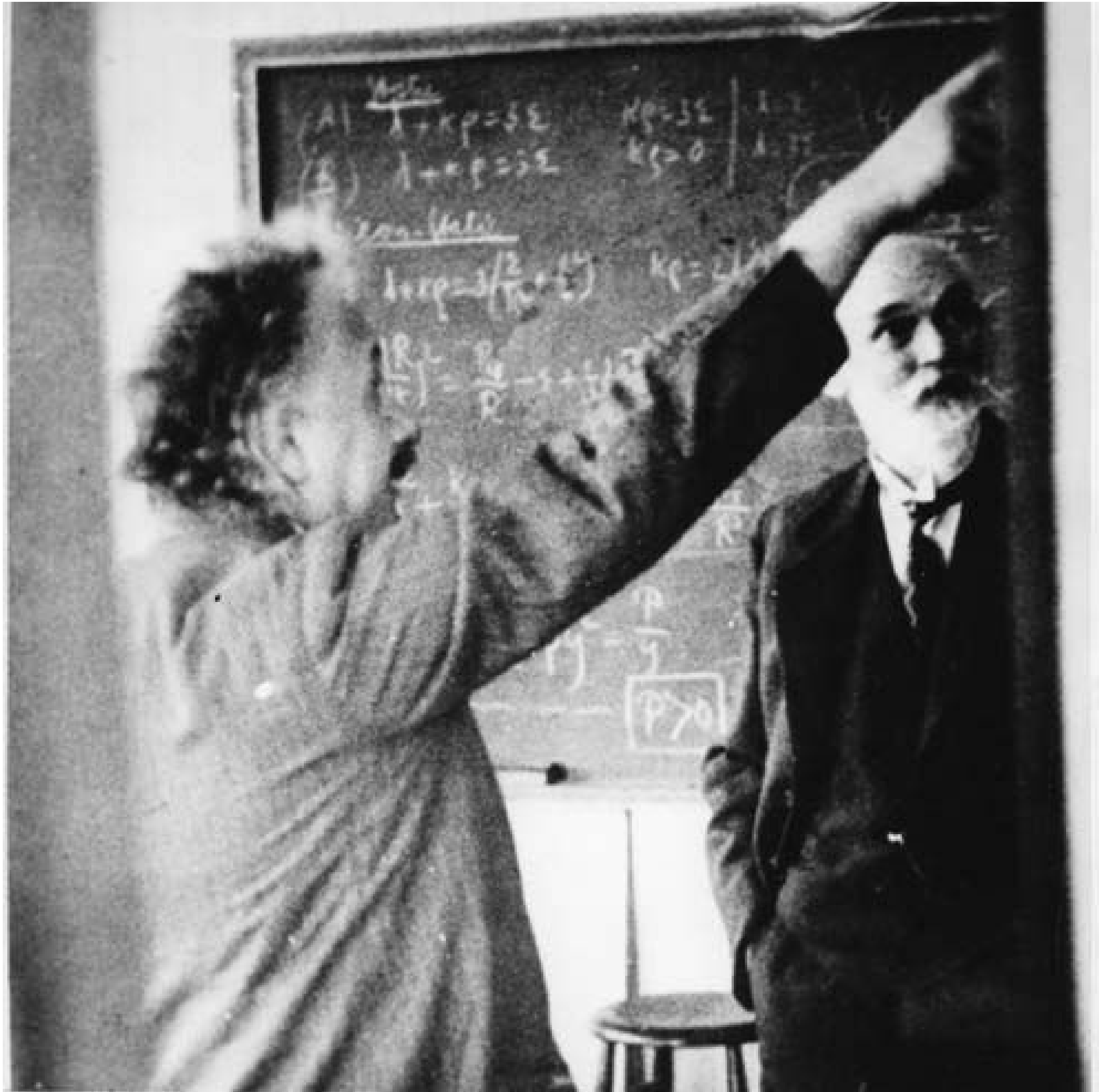


图 3.15 1932年，爱因斯坦和德希特正在帕萨迪纳的加州理工学院研究他们的宇宙模型

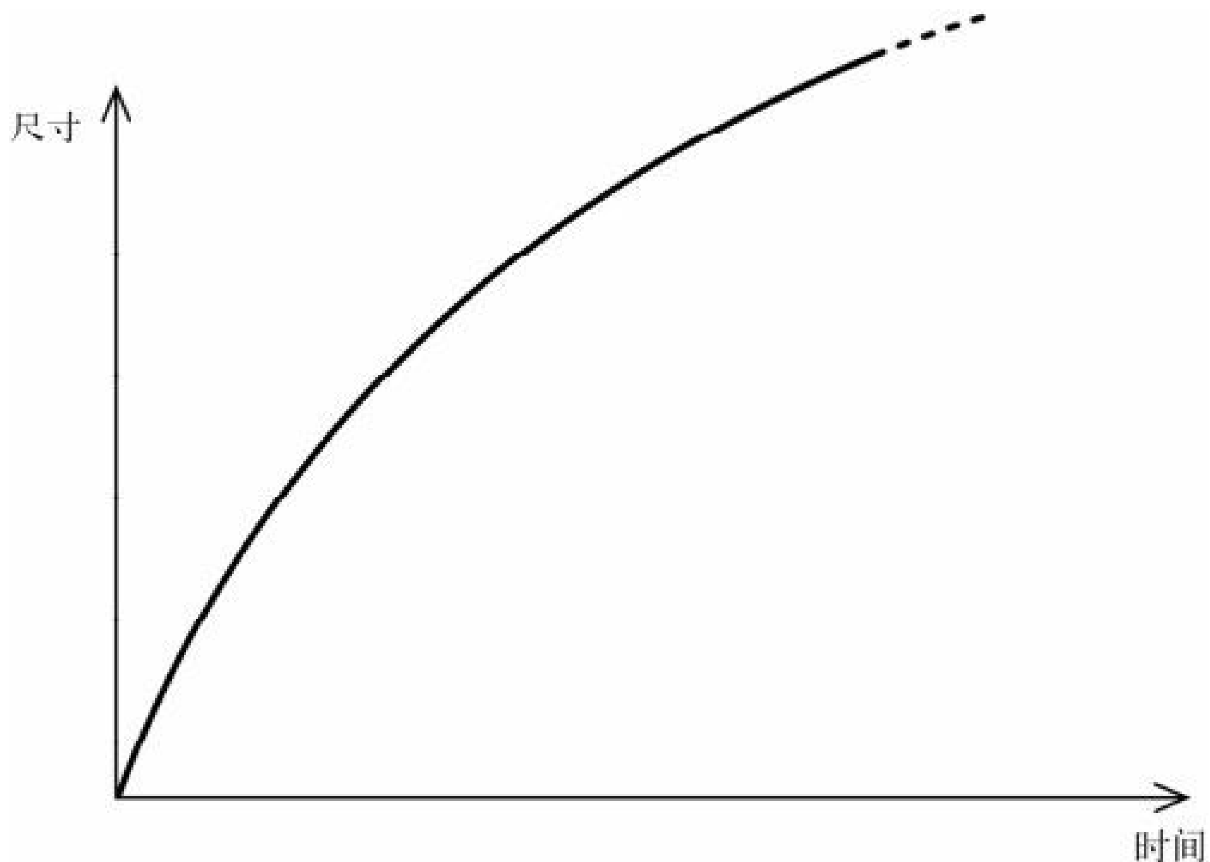


图3.16 爱因斯坦–德希特宇宙，距离增大与时间的2/3次方呈正比

所有这些宇宙模型都有一个简单的、带有启发性的解释，想必牛顿都很好理解。如果你朝天上扔石头，石头会落回大地，因为石头没有足够的能量克服地球产生的引力。但如果你以超过每秒11公里的速度把它扔出去，它就不会回来了。[42]宇宙也有一个类似的“逃逸速度”，决定了它是否能永远膨胀下去。闭合的、最终会塌缩的宇宙对应于膨胀速度小于逃逸速度的情况；开放的、负曲率的宇宙的膨胀速度大于逃逸速度。爱因斯坦–德希特宇宙，曲率为零，膨胀速度严格地与逃逸速度相等，刚好可以永远膨胀。无论其中包含物质的密度增加多少，无论多么微小，都会导致膨胀速度小于它的“逃逸速度”，将来会开始塌缩。

托尔曼的振荡宇宙

朝上的，必落下来。

——弗雷德里克·波特尔（1897～1987，美国学者）[43]

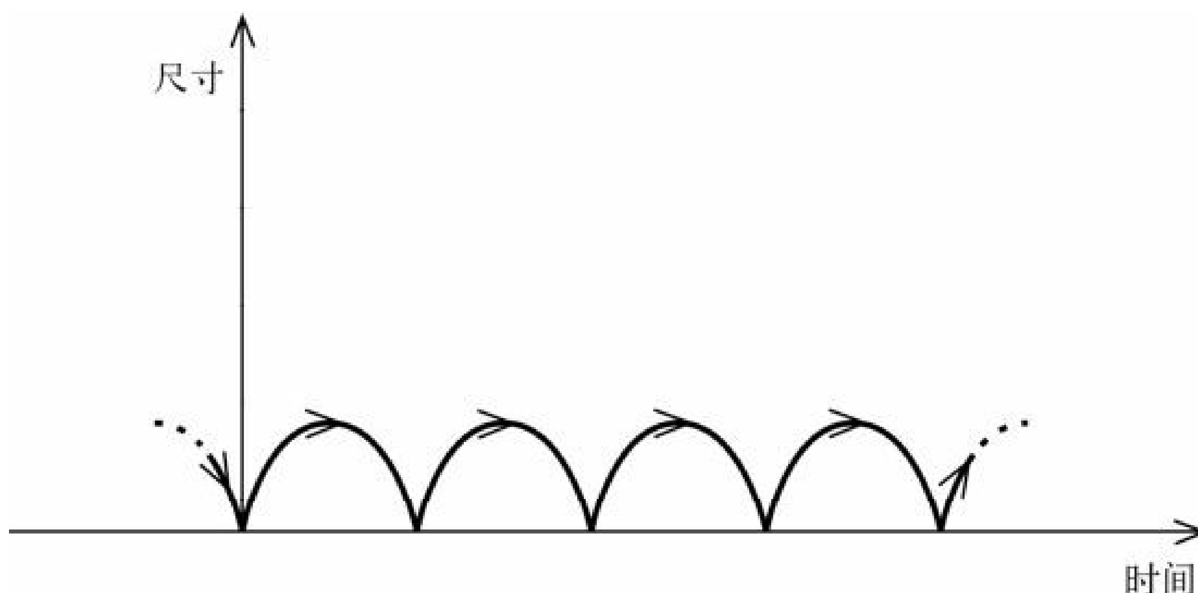


图3.17 周期性振荡宇宙。每次都先是一场大塌缩，然后是开始新一轮的膨胀，依此类推直至永远

弗里德曼的开创性研究第一次提出了体积膨胀到最大后又收缩到零的宇宙（图 3.7）。他暗示还有一种可能性是，存在一系列这样首尾相接的循环，宇宙一次又一次地达到最大体积和最小的零体积，像一个反弹的球。就像我们强调过的那样，弗里德曼对天文学的事不太感兴趣，他只是想从数学的角度找到一些爱因斯坦方程组的解。然而从宇宙学的角度讲，一个闭合的宇宙，从过去到未来永远在振荡，这其中会有些有趣的问题。会如图3.17所示，所有的循环都是一样的吗？

1932年，理查德·托尔曼（Richard Tolman，1881～1948）开始考虑这个问题。他在帕萨迪纳的加州理工学院工作，这也是爱因斯坦经常访问的学术机构。托尔曼的学术背景和其他宇宙学的开拓者大不相同。他是加州理工学院物理化学和数学物理专业的教授，还对热力学有独特的兴趣。他想，如果在爱因斯坦方程组的无限循环振荡宇宙的解中，应用著名的热力学第二定律会怎样。每一个循环里的熵都较前一个有所增加，就像从有序的能量，如原子和星系那样的，转化成无序的热辐射。向爱因斯坦的宇宙学方程组中引入一个转化过程很简单，这意味着振荡宇宙的每一个循环中辐射能量（有正的压强）所占比例都比前一个大。这使得每一个循环的最大尺寸都比前一个的大（图3.19）。振荡的尺寸越来越大，持续时间越来越长，宇宙最终看起来就像爱因斯坦–德希特宇宙，因为膨胀的时间越来越长。

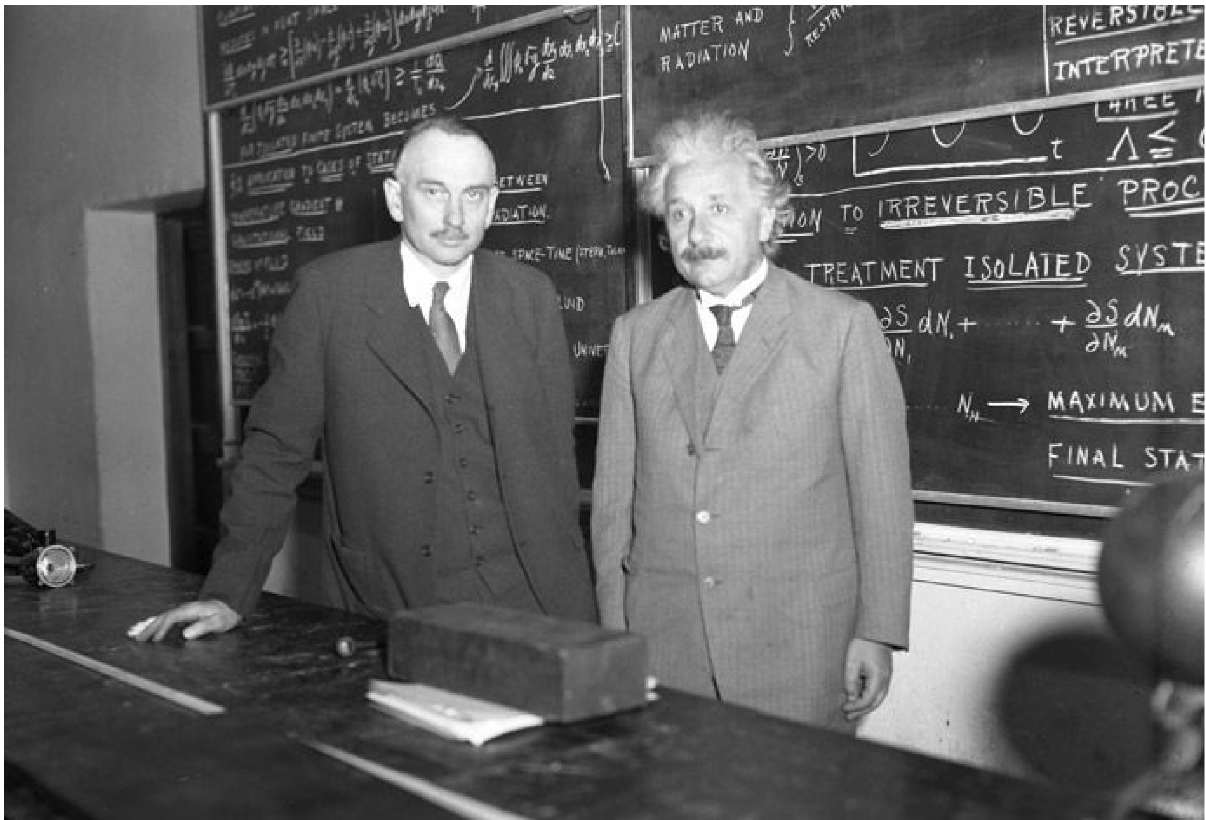


图3.18 理查德·托尔曼在一次学术演讲后，为爱因斯坦解释他的振荡宇宙模型中的热力学机制。他们身后黑板上的方程描述了膨胀宇宙中物质和辐射的热力学性质

如果我们逆着时间往回看这种宇宙，也许不需要一个开端。最后，如果我们跳回很久很久的以前，这种宇宙一定非常小，以至于量子效应主导了引力的行为，爱因斯坦方程组就不再成立了。也可以说，每次宇宙一路收缩到零的时候，爱因斯坦方程组也不成立，因为这时物质和辐射的密度是无限大。然而，如果有新的量子引力机制使得宇宙在很小但不是零的尺寸下发生反弹，不断增长的循环振荡还是有可能的。

托尔曼在科学上很谨慎，他对特殊宇宙模型的许多研究其实意在警示那些对宇宙本质贸然得出结论的冒失鬼。因此，他从尺寸和熵不断变大的反弹宇宙模型中得到的结论简单如下：

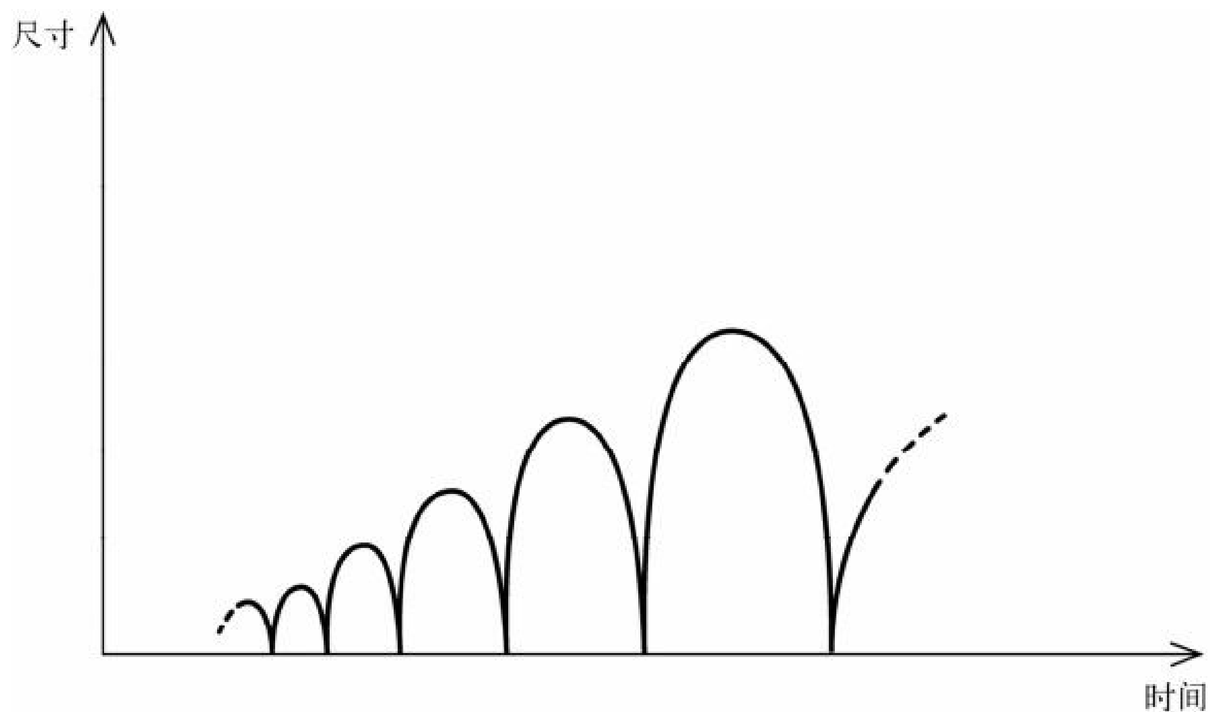


图3.19 托尔曼的振荡宇宙。能量守恒和宇宙熵的增加使得宇宙的尺寸随着循环次数的增加而越来越大

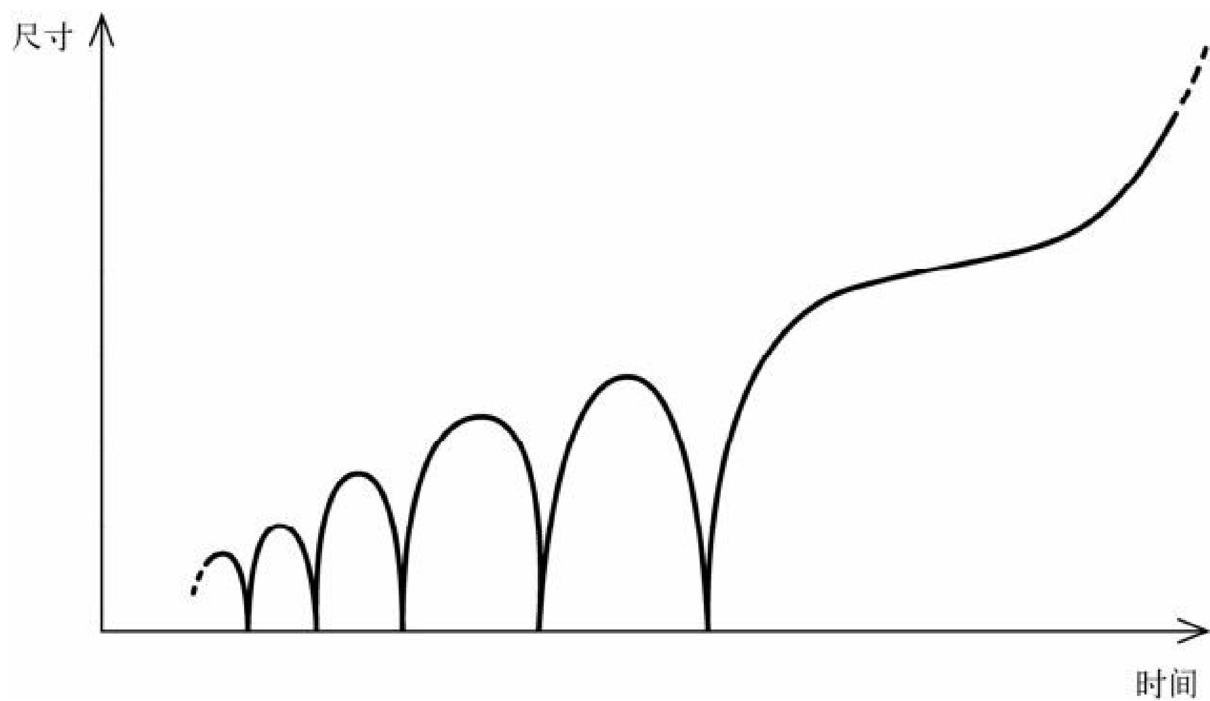


图3.20 巴罗和东布罗夫斯基的宇宙模型，其中有一个正的拉姆达力。无论拉姆达的取值有多么小，托尔曼宇宙的振荡最终都会停下来，并永远

膨胀下去，越来越像德希特宇宙

如果我们不再武断地认为，热力学的原则必然要求一个宇宙从无限的过去开始，并注定在将来走向衰败和死亡，至少这看起来是最明智的。[44]

但托尔曼忽略了反弹宇宙的一个有趣细节，直到很久以后的1995年才被马留斯·东布罗夫斯基（Mariusz Dąbrowski）和我发现。[45]如果宇宙中存在爱因斯坦排斥性的宇宙学常数，无论它的数值多么小，最终都会导致振荡结束，宇宙膨胀开始加速，就像德希特宇宙那样，永远不再收缩和振荡（图3.20）。[46]

勒梅特和托尔曼的古怪宇宙

但谁又能在无限膨胀的马鞍/上安心跨鞍？

——W. H.奥登（1907～1973，美国诗人）

1933年，勒梅特和托尔曼同时开始思考，如何才能找到一个比先前发现的更加真实的宇宙模型。在爱因斯坦的指引下，所有早期寻找爱因斯坦方程组所描述的宇宙的过程，都使用了同一个简单的假设，那就是宇宙的每个地方和每个方向都是一样的。但真正的宇宙不是这样子的。如果真是那样的话，你就不可能存在了。而且，人们还希望宇宙对完全对称的偏离程度很小，可以忽略不计，至少刚开始描述真实的宇宙时可以忽略。然而，当你朝宇宙中望去时，你会发现恒星和星系在天空中的分布极为不规则。我们可以找到这样的爱因斯坦方程组的解吗？

勒梅特和托尔曼都成功了，他们用不同的方法找到了第一个不均匀的宇宙；这类宇宙的特点是，物质密度和膨胀速率会随着时间和地点的变化而变化。勒梅特的想法非常富有远见。他想知道这是否能帮助我们理解为什么宇宙中形成了像恒星、星系这样的结团。尤其是他能算出这些高密度区域的演化行为，发现随着时间的推移，它们的结团会越来越明显（因为它们的引力会吸引更多的物质，使得空旷的地方更空旷）。[47]因此，宇宙的结团逐渐加剧，并且会越来越严重。

托尔曼没有勒梅特那么强烈的冒险精神，他很担心从无法证实的假设出发得到的结论是否可靠，尤其当时所有关于爱因斯坦方程组的研究都假设宇宙各处是均匀的。他认为这种假设不太可能是真的，因此，由此得出的宇宙开端或者宇宙未来终结于大塌缩的结论过于冒险，不靠谱。这些结论依赖于一种纯粹的假设，那就是宇宙中我们看不到的部分和我们看得到的部分是一样的。

当托尔曼找到爱因斯坦方程组的不均匀宇宙的解时，他发现这些宇宙的不同区域可以表现得像不同类型的弗里德曼–勒梅特宇宙。不同区域的空间曲率和物质密度可以连续变化。某个区域中的物质密度可以比平均密度大，空间曲率为正。这部分区域就像一个闭合的宇宙，在自身引力的作用下，尺寸膨胀到一个极大值之后，再塌缩回来，也许形成了

一个星系。在一个比平均值空旷的区域，空间曲率为负，而那里的物质产生的引力永远无法抵抗膨胀，无法开始塌缩。托尔曼认为，这使得我们可以设想这样一种可能性：

宇宙中我们望远镜鞭长莫及的地方也许在收缩而不是膨胀，其中的物质密度、物质演化的阶段也与我们所熟悉的情况非常不同。明智的做法是，对于期望中的宇宙初始状态，不要根据均匀模型中得到的现象而做出太肯定的结论。[48]

他做出不要对宇宙的开端草率下结论的警告，是因为爱因斯坦方程组允许的这种古怪宇宙有三个特点。首先，宇宙处处各不相同意味着宇宙各处的膨胀没必要同时开始。其次，人们意识到可能存在一种极端情况，即宇宙某些区域的膨胀开始得特别晚，以至于膨胀早的区域可以看到它们，影响到它们。[49]最后，考虑存在宇宙学常数的情况，也许有些区域不存在开端（因为这些区域可以局域地看作爱丁顿–勒梅特宇宙），而其他区域却有一个开端。

对于宇宙的最终状态以及宇宙究竟是无限的还是有限的，宇宙各处也会存在区别。托尔曼指出，我们可能生活在无限大宇宙的一个巨大的、致密的、正曲率的区域中，这个宇宙会永远膨胀下去，而我们可见的区域却可以像一个有限闭合的宇宙一样，最终将开始收缩。如果假设整个宇宙的性质和我们附近的一小块地方一样，那么我们就可能做出关于宇宙终极特性的错误结论。

米尔尼的宇宙（和牛顿的宇宙）

我们没有探测到任何膨胀效应或退行因子。已有的数据仍然支持静态宇宙，而不是快速膨胀的宇宙。

——埃德温·哈勃（1889～1953）[50]

20 世纪 30 年代中叶，人们对简单的膨胀宇宙模型有了充分的理解。埃德温·哈勃积累了大量红移观测的数据，并强有力地证明了遥远星系的退行速度和它们到我们的距离呈正比。他验证了勒梅特在1927年第一次导出的公式，但他看起来不愿意承认自己的数据说明宇宙在膨胀。人们也深刻地理解了爱因斯坦方程组简单的、各向同性、均匀的宇宙解。英国宇宙学家威廉·麦克里和阿瑟·米尔尼（Arthur Milne，1896～1950）通过自己的工作，将这样的认识在没有那么擅长数学的天文学家中传开了。他们证明了，通过求解艰深的爱因斯坦方程组才得到的所有宇宙演化行为，其实都可以不必解方程组，而直接从牛顿的旧引力理论中得出。[51]他们描述了在牛顿著名的平方反比引力的作用下，一团球形物质会如何演化。如果你想算上爱因斯坦宇宙学常数的排斥作用，只需直接在两个物体的牛顿平方反比定律中增加一项拉姆达力。如果物质粒子的动能大于引力势能，宇宙就会永远膨胀。但如果引力势能更大，那么膨胀就会到达一个极大值，然后开始收缩。两者之间的情况是爱因斯坦–德希特宇宙——膨胀的能量和引力势能完全相等，就像地球上发射了一个火箭，速度刚好达到“逃逸速度”。这种看待爱因斯坦最简单的宇

宙的方法，仍然是给物理系和天文系学生上课时用到的最通俗的方法。

米尔尼是一个年轻而杰出的天体物理学家，他的主要贡献在于恒星和恒星大气的有关理论。他获得了剑桥大学三一学院的入学奖学金，是史上所有申请者中的最高分。他二十多岁时就成为皇家学会的会员，1928年就当上了牛津大学数学系的教授。自1932年起，他转而研究宇宙学，发展了一种相对论运动的独特理论，并挑战了人们广泛接受的宇宙热寂的观念。

米尔尼证明，不借助爱因斯坦的引力理论，也能研究哈勃定律和膨胀宇宙。他开始从不同角度讨论原子经历的“时间”是否有必要和“引力”系统经历的时间相同。但是这个他所喜欢的新的膨胀宇宙模型，用爱因斯坦方程组描述起来也非常简单。米尔尼的宇宙有些像德希特宇宙，不包含物质，但不同的是，它也不包含宇宙学常数的排斥力。它只有负的空间曲率，永远以光速膨胀（图3.21）。[52]

米尔尼的宇宙是满足各向同性均匀假设的可能宇宙中最简单的一个。尽管这个宇宙听起来很特殊，不切实际（真实的宇宙毕竟不是空无一物、没有物质和辐射的），但仍然值得关注，因为如果你往一个不含宇宙学常数的开放宇宙中加入普通物质和辐射，它会永远膨胀，越来越接近米尔尼宇宙的样子。[53]

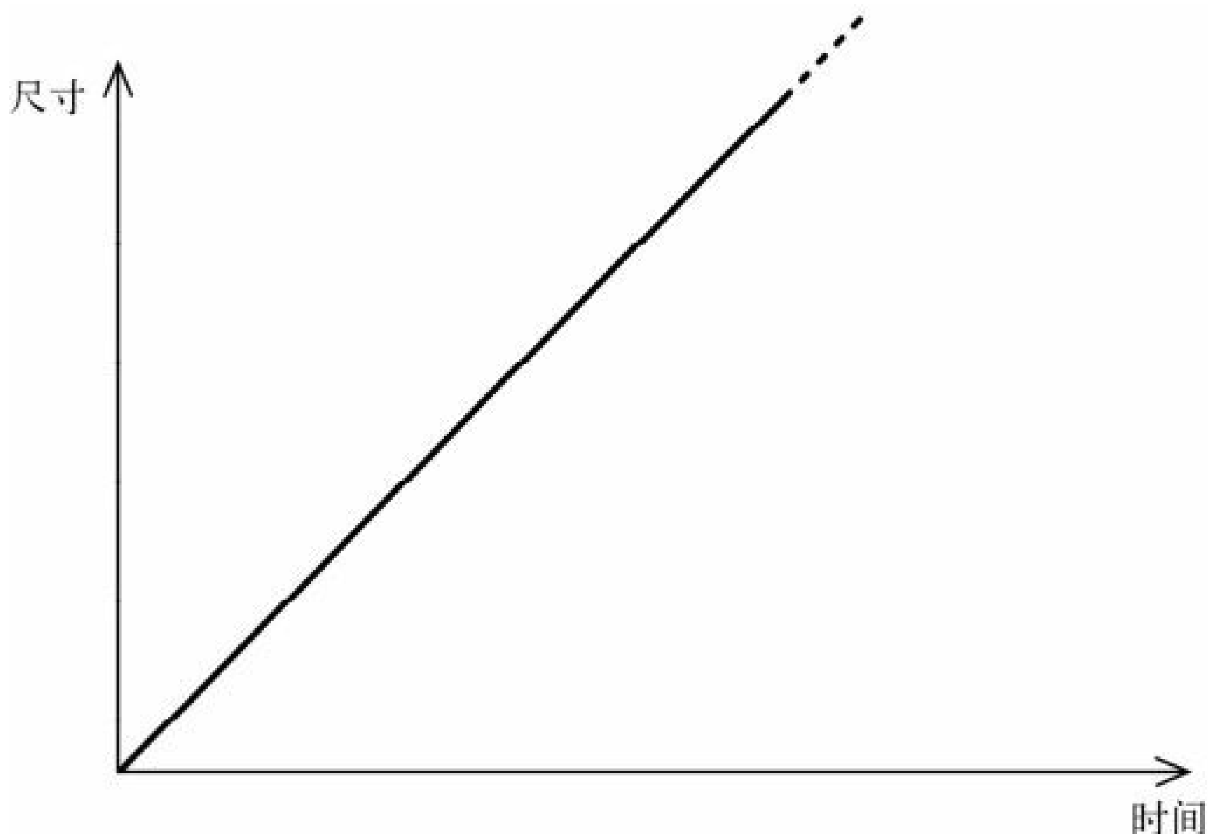


图3.21 米尔尼的宇宙，距离与宇宙年龄呈正比

身为科学家，米尔尼还拥有强烈的（圣公会）宗教信仰和动机，想要研究如何将宇宙膨胀的科学理论同基督教的上帝和神创论统一起来。他在这方面进行了大量的写作，尤其重要的是1952年在他去世[54]后出版的《现代宇宙学和基督教的上帝》一书。[55]不过，他从没有打算写给公众看，他的书只有物理学家和数学家才能读得懂。他非常喜欢无限宇宙，因为这给上帝开展无数演化实验留下了广阔空间。他认为有限的宇宙留给神的空间很有限。

米尔尼的宇宙学想要推导出宇宙几何可能存在的“合理”结构，他相信这种合理结构对上帝创造的宇宙种类施加了一些限制。他相信，宇宙的不同方向不可能以不同的速率膨胀——尽管我们在下一章会看到，其他人不这样认为。它必须有一种全局的法则——到处都一样。他坚持认为自然法则由宇宙的结构合理地决定：引力的定律不可能既发生轻微的改变，同时又不破坏整个宇宙的和谐性。所以对米尔尼来说，逻辑从某种程度上限制了上帝的能力，它也决定了宇宙及其法则可能拥有的形式。[56]然而在某种程度上，这一点已经被无限空间所引入的自由度平衡了：米尔尼相信，从无限的宇宙中能够发掘出生命的无限可能性。有趣的是，这和勒梅特在此问题上的看法完全相反。勒梅特强烈地倾向于宇宙的有限性。对他来说，宇宙存在的目的就是让生物去探索和感悟，而无限的宇宙不允许那样：它超越了人类的理解范围，任何事都有可能在某处发生；对其中的居民来说，它是个过于广阔的家园。

注释

[1] R. W. Clark, *Einstein: The Life and Times,* World Public. Co., New York (1971), pp.385–386.

[2] 施特劳给英国前总检察长戈德史密斯（Lord Goldsmith）的信，引自 2010年1月26日BBC网站：http://www.bbc.co.uk.

[3] J. Eisenstaedt, *The Curious History of Relativity,* Princeton University Press, Princeton (2006), pp. 123–124.

[4] 我们已经观测到像地球这样的旋转物体的周围存在引力的拖曳效应。质量和运动所产生的引力不仅改变了空间的形状，还改变了时间流逝的速率。在强引力场中，空间高度弯曲，时钟的“滴答”也会比在引力较弱的空间较平坦的地方的时钟慢一些。这个效应经常在实验中观察到。

[5] 当时见解最为深刻的人是埃里希·克莱舒曼（1887～1973）。他证明，爱因斯坦理论中的“广义协变”性质，在别的用张量语言所描述的理论中同样存在。参见：E. Kretschmann, Annalen der. Physik 53, 575

(1917).关于克莱舒曼的论文和爱因斯坦的回应，详细内容请见：R. Rynasiewicz, ‘Kretschmann’s Analysis of Covariance and Relativity Principles’, in H. Goenner, J. Renn, J. Ritter and T. Sauer (eds.), *The Expanding Worlds of General Relativity*, Birkhäuser, Boston (1999), pp. 431–462.

[6] J. Church,A Corpse in the Koryo, Thomas Dunne, New York (2006), p. 266.

[7] 量子力学为一切有质量的物体都赋予了波动的性质。波长和物体的质量呈反比。质量小的物体，波长就会比较大。如果量子的波长比物体的物理尺寸还大，物体的量子性质就会非常明显。但是如果物体的质量非常大，就像你我这样的庞然大物时，物体的量子波长就会远远小于物理尺寸。只要运动得不太快，它就会符合牛顿的预言。

[8] N. Coward,*Design for Living*, Act 3, Scene 1 (1933).

[9] 对此，有人曾提议对牛顿定律略微进行改动，也有人曾考虑太阳并不是完美的球体。

[10] A. Einstein, *Collected Papers of Albert Einstein*, vol. 6, p. 21.这句话最早是在1915年说的。

[11]也可以说每一点都是球面的中心。

[12]如果你问，什么样的引力定律能够保证一个球体产生的引力等价于一个位于球心的相同质量的质点产生的引力，答案是正比于 $1/r^2$或 r 的引力定律。这种定律的一般形式可以写成两者的线性叠加$-A/r^2+Br$，其中 A和 B是常数。当 $r^3=A/B$时，合力和加速度就等于零了。在引力相对较弱时，爱因斯坦更为复杂的数学理论也能给出类似的效应。这里的常数B就等于所谓的宇宙学常数乘以光速的平方。

[13]即$t=2\pi R/c$，R是空间的半径，c是光速。

[14]如果物质的平均密度是 ρ克每立方厘米，那么光波环绕宇宙一周的时间就是$2/\sqrt{\rho}$小时。用宇宙的质量M、光速c和牛顿常数G表示爱因斯坦宇宙的半径，是$R=2GM/\pi c^2$。注意，这跟史瓦西黑洞半径$R=2GM/c^2$相差一个系数π，这是因为爱因斯坦的静态宇宙的空间几何是球对称的非欧几何。于是，宇宙的体积就应该是$2\pi R^3$而不是$4\pi R^3/3$。

[15]爱因斯坦和德希特之间的通信内容非常丰富，十分精彩。米歇尔·詹森（M. Janssen）的一篇综述写得非常好，其中还有作者对通信内容的评论，参见：http://www.lorentz. leidenuniv.nl/~vanbaal/ART/E-dS.pdf.

[16]他把爱因斯坦的静态宇宙称为“A解”，为了进行区分，把自己的解称为“B解”。

[17] A. Eden, *The Search for Christian Doppler*, Springer, New York.

(1992).这本书把多普勒的开创性研究《论双星和天上其他恒星的颜色》（*On the Coloured Light of the Binary Stars and some Other Stars of the Heavens*）翻译成了英语。1845年，荷兰物理学家克里斯托弗勒斯·柏斯·巴洛特（Christophorus Buys Ballot，1817～1890）请了一个小型管弦乐队在乌德勒支开往阿姆斯特丹的火车上演奏标准音，通过实验验证了声音的频率和声源速度之间的关系。

[18]弗里德曼的名字有很多种不同的拼法：Friedmann、Friedman、Fridman 等。我们将采用最常见的拼法：Friedmann。但请注意，1922年他在德国《物理学杂志》（*Zeitschrift für Physik*）上发表第一篇文章时署名A. Friedman，1924年第二篇文章上的署名是A. Friedmann。而他在图书封面的署名以及他的出生记录，从俄文逐字翻译过来都是Aleksandr Aleksandrovich Fridman。

[19]他的工作是对炮弹的落点作出预测，有时他甚至要在炮火攻击时乘坐飞机来评估自己预测的精确度。由于工作出色，他还为此获得了勋章。

[20]乔治·伽莫夫在自传《我的世界线》（My World Line）里说，是一次热气球的高空旅行导致了弗里德曼的死亡，但是没有证据能够证明这一说法。伽莫夫当时是圣彼得堡大学的一个学生，一直希望成为弗里德曼的研究生，但弗里德曼的去世使他的希望破灭了。关于弗里德曼的传记，参见：E. Tropp, V. Frenkel and A. D. Chernin,Alexander A. Friedmann: *The Man Who Made the Universe Expand*, Cambridge University Press, Cambridge (1993).

[21]这个数值跟现在的观测结果（宇宙的年龄是137亿年）非常吻合，我们还不清楚弗里德曼为什么用了这个数值，也许只不过是随便猜的。他说这个数值“只能作为相关计算的一个例子”。

[22] H. Nussbaumer and L. Bieri,*Discovering the Expanding Universe*, Cambridge University Press, Cambridge (2009), p. 90.弗里德曼的书现在还只有德文版。

[23]在书中，弗里德曼提出了一个极为深刻的观点。他指出，尽管正曲率的球对称空间的体积是有限的，但是反过来说就不一定了。负曲率的开放空间的体积可以是无限大，也可以不是。这取决于某种无法由爱因斯坦方程组决定的性质：空间的拓扑性质。我们将在第11章对此进一步讨论。弗里德曼也将爱因斯坦的宇宙学常数引入了他的方程组中。

[24]弗里德曼两篇论文的英文版，“论空间的曲率”（‘On the Curvature of Space’）和“论一个世界的空间曲率为负常数的可能性”（‘On the Possibility of a World with Constant Negative Curvature of Space’），以及爱因斯坦发表的两个评论，可以参见：J. Bernstein and

G. Feinberg,*Cosmological Constants*, University of Columbia Press, New York (1989), pp. 49–67;*General Relativity and Gravitation* 31, 1991–2000 and 31, 2001–2008 (1999).

[25] V. A. Fock,*Soviet Physics Uspekhi* 6, 414,转引自：H. Kragh,*Cosmology and Controversy*, Princeton University Press, Princeton, p. 27.

[26] G. Lemaître,*The Primeval Atom: An Essay on Cosmogony*, transl. B. H. and S. A. Korff, Van Nostrand, New York (1950).

[27]这原是天主教学生的宿舍，后来成为剑桥大学的圣埃德蒙学院。

[28] G. Lemaître, ‘Un univers homogène de masse constante et de rayon croissant rendant compte de la vitesse radiale des nébuleuses extragalactiques’, (‘A Homogeneous Universe of Constant Mass and Growing Radius Accounting for the Radial Velocity of Extragalactic Nebulae’),*Annales de la Société Scientifique de Bruxelles, série A* 47, 49 (1927).

[29]我在苏塞克斯大学的同事、已故的威廉·麦克里爵士，对宇宙学研究早期的大部分领军人物都很熟。有一次他跟我说，他总认为勒梅特是这些科学家之中（除了爱因斯坦以外）最厉害的一个，他总能用最简单的方法得出重要结果。

[30]据此得出的哈勃常数是575千米每秒每百万秒差距（km/sec/Mpc），但这个计算没有在1931年论文的英文版翻译中提到。

[31] 1930年哈勃写给德希特的信，引自：Nussbaumer and Bieri,*Discovering the Expanding Universe*, pp. 130–131.

[32] A. S. Eddington, ‘On the Instability of Einstein’s Spherical World’, *Mon. Not. R. Astron. Soc.* 90, 668 (1930).那段时间，爱因斯坦正好去剑桥拜访了爱丁顿，所以我们可以设想爱因斯坦或许曾从该论文作者那里亲口得知了这一问题。

[33] G. Lemaître,*Mon. Not. R. Astron. Soc.* 91, 483 (1931).

[34]爱因斯坦在放弃了静态宇宙的概念之后，便抛弃了宇宙学常数，并把引入宇宙学常数一事称为“一生中最大的错误”。但是其他宇宙学家，如爱丁顿、德希特和勒梅特都认为宇宙学常数是这一理论的重要内容。勒梅特进一步证明，如果你不承认理论上存在这种新的引力形式的话，也可以把宇宙学常数理解为宇宙的一种真空能（vacuum energy）。爱丁顿把宇宙学常数看作是引力理论与量子理论和研究质子、电子的粒子理论相联系的纽带。他断言，宇宙学常数是爱因斯坦理论中最重要的组成部分。在最近一些年里，宇宙学常数已经从理论假设

变成了实验事实，我们将在第12章进一步讨论。

[35]你可以简单地认为宇宙一直是静止的，然后从过去某个有限的时刻开始膨胀。爱丁顿–勒梅特宇宙是闭合的，在无穷远的过去是静止的，此后一直在膨胀。

[36] A. S. Eddington,*The Expanding Universe*, Cambridge University Press, Cambridge, p. 56.

[37]关于相关问题的讨论，参见：O. Godart and M. Heller, *Pont. Acad. delle Scienze, Commentarii* 3, 1–12, 1929,转引自：H. Kragh,*Cosmology and Controversy*；还可参见：O. Godart and M. Heller,*Cosmology of Lemaître*, Pachart Publ., Tucson (1985).奥登·歌达特（Odon Godart）是勒梅特的科研助手，而迈克尔·海勒（Michael Heller）像勒梅特一样，是天主教牧师兼宇宙学理论家。具有讽刺意味的是，尽管勒梅特不愿意从自己的宇宙学理论出发讨论任何形式的自然神学或者宗教辨惑学问题，但他发现有人却敢于借题发挥。1951年，罗马教皇庇护十二世（Pope Pius XII）在勒梅特任院长的教皇科学院进行了一场关于宇宙创生的著名演说（不是勒梅特起草的）。教皇把勒梅特的宇宙从过去某个时刻开始膨胀的科学理论，类比于无所不能的神“从无中创造”，将它看作是这个古老教条的现代版本。勒梅特在短暂的余生中试图温和地掐灭这股自以为是的火苗，但没有取得什么明显的效果。

[38] E. R. Harrison,*Mon. Not. R. Astron. Soc.* 137, 69 (1967).

[39]这两句关于合作论文的有趣评论，分别是爱因斯坦跟爱丁顿的讨论以及德希特给爱丁顿的信中说的。1936年，爱丁顿在剑桥的一次演讲中提到了这两句话，演讲的文字参见：‘Forty Years of Astronomy’in *Background to Modern Science*, Cambridge University Press, Cambridge (1940).

[40] A. Einstein and W. de Sitter,*Proc. Nat. Acad. Sciences* 18, 213 (1932).

[41]其中的距离总是随着时间的2/3次方增加，即$R \propto t^{2/3}$。

[42]这一临界速度，也叫做地球的逃逸速度，等于 2GM R ，其中M 是地球的质量、R是地球的半径。

[43]这句谚语说的是牛顿的引力定律，第一次见于字面是在：F. A. Pottle,*The Stretchers: The Story of a Hospital Unit on the Western Front*, Yale University Press, New Haven (1929).不过很可能已经口口相传了至少五十年。

[44] R. Tolman, *Relativity, Thermodynamics and Cosmology*, Clarendon Press, Oxford (1934), p. 444.

[45] J .D. Barrow and M. Dąbrowski, ‘Oscillating Universes’,*Mon. Not. R. Astr. Soc.* 275, 850(1995).

[46]当宇宙变大到一定程度以后，宇宙学常数的排斥作用就会开始占优，胜过其他物质产生的引力。熵的增加使得宇宙每经历一个周期就变大一些，直到宇宙学常数接管了宇宙的膨胀。此时，膨胀永远不会停下来，即使空间是闭合的，也会永远膨胀下去，就像勒梅特宇宙一样。

[47] G. Lemaître, ‘The Expanding Universe’,*Ann. de la Soc. Scientifique de Bruxelles A* 53, 51 (1933).英文翻译参见： M. A. H. MacCallum, *Gen. Rel. Gravitation* 29, 64l (1997).

[48] R. Tolman, *Proc. Natl. Acad. Sci USA* 20, 169 (1934).

[49]在这种极端的情况下，我们把初始时刻时宇宙的空间截面称为“类时的”（timelike）。这就表明此时的空间中，有些地方处于另一些地方的因果未来（causal future）。如果这种事情不可能发生，我们就把初始截面称为“类空的”（spacelike）。

[50] E. Hubble, ‘Problems of Nebular Research’, *Scientific Monthly* 51, 391–408 (November 1940),该段话在第 407页；转引自：H. Kragh,*Matter and Spirit in the Universe: Scientific and Religious Preludes to Modern cosmology*, Imperial College Press, London (2004), p. 153.

[51] E. A. Milne,*Z. f. Astrophysik* 6, 29 (1933), reprinted in *Gen. Rel. Gravitation* 32, 1939 (2000); W. H. McCrea and E. A. Milne,*Quart. J. Math.Oxford* 5, 73 (1934), reprinted in *Gen. Rel. Gravitation* 32, 1949 (2000).

[52]宇宙的膨胀随时间的变化为R=ct。

[53]随着时间的流逝，随着宇宙的膨胀，物质和辐射产生的引力比负曲率的引力效应衰减得更快，最终，宇宙的行为会变得像米尔尼的特殊宇宙模型一样。例如，如果往米尔尼的宇宙中加入黑体辐射，宇宙尺寸的变化率就是$R^2=t-kt^2$，其中k是曲率。如果k是负的，那么当t非常大时，R就和t呈正比，就很像米尔尼的宇宙模型。如果t很小，R就正比于t的平方根。注意，当 k为正的时候（例如，令 k= +1），宇宙是闭合的，此时R(t)的变化曲线是一个半圆形，从 t= 0时出发，在 t= 1/2时达到最大值，然后终结于 t= 1的时刻。

[54] 1950年 9月，米尔尼年仅 54岁就因心脏病去世了，当时他在都柏林参加皇家天文学会的一次会议。

[55] E. A. Milne, *Modern Cosmology and the Christian Idea of God*, Clarendon, Oxford (1952).他早前出的一本书也有关于这个问题的一段内容，参见：*Relativity, Gravitation and World Structure,* Clarendon, Oxford (1935), p. 138.他对此的观点比通常预计的更为高深。由于他的宇

宙模型不会遭遇热寂，所以他认为，既然宇宙的年龄、起源、膨胀速度和尺寸都跟计时系统的选取有关，因此宇宙的性质必然跟具体观测者相关。例如，有些观测者会发现引力常数G会随时间变化，而有些观测者就不这么认为。米尔尼的模型跟实验观测的关系不大，只是依赖于空间和时间的均匀性，而不像勒梅特的宇宙那样，将理论和观测有机地结合在了一起。

[56]在米尔尼看来，自然法则的形式也是一种必然真理（necessary truth），就像欧几里得几何中的三角形内角和等于180度一样。

第4章
出乎意料的宇宙

就像在一个滑雪胜地里，到处是钓老公的姑娘，到处是钓姑娘的老公，这事并不像表面看上去那样对称。

——阿兰·林赛·麦凯（Alan Lindsay Mackay，英国晶体学者）

分形的宇宙

考虑宇宙的伟大或无限性没有意义，这就好像对于瑞丁监狱的一名囚犯来说，他会很高兴地听闻监狱已经覆盖了全国一半的地方。

——吉尔伯特·K.切斯特顿（1874～1936，英国作家）[1]

20世纪 20年代，爱因斯坦的理论导致了一门现代学科的诞生，即从整体上研究宇宙的宇宙学。确实，不同的宇宙模型变得像12路公交车的不同班次一样，没什么稀奇的。然而与此同时，仍有一些科学家负隅顽抗，试图说服爱因斯坦，用牛顿的旧理论来描述宇宙并不像爱因斯坦说的那样不可能。

但他们面对的主要困难在于，应该如何处理物质均匀分布的无限大空间。要么宇宙中所有的物质必须被限制在一个有限质量的大型岛屿之中，要么牛顿描述引力如何随距离增大而逐渐减弱的著名的平方反比定律必然在某个巨大尺度上失效。在一个密度均匀的无限大宇宙中，你可以“证明”，你从宇宙中所有物质感受到的引力可以等于任何值！[2]明显有地方出错了。

19世纪末的许多宇宙学模型解决了这个佯谬，如前面我们见到过的有球状拓扑的史瓦西宇宙。1907 年，爱尔兰科学家埃德蒙·富尼埃·达尔伯（1868～1933）面向公众出版了一本有趣的小册子，题为《两个新世界》。他在其中提出，可以自然地认为宇宙有一套分层结构，涵盖从原子到太阳系以至更大的体系，无穷无尽。[3]实际上，关于宇宙是一个大型星系群的一部分，而后者又属于更大型的星系群，以至无穷的说法，我们早在说到赖特、康德和兰伯特的研究时就已见过了。伟大的美国作家埃德加·爱伦·坡（1809～1849）也描述了这样一种印象。在1848年完成的散文诗《我发现了》中，他提到了由“星系群的星系群”的无穷序列构成的宇宙，每一层结构都有不同的法则，和人类没有任何联系。[4]

富尼埃·达尔伯提出这个想法的一个动机是，他想给出一种系统阐述质量无穷大宇宙的模型的方案。尽管他的设想不合常规，但他本人还是受到了严肃对待。他是电学和磁学方面的专家，并在1923年主持传输了世界上第一幅电视画面（画面的内容是英王乔治五世）；他还发明了光声机，将光信号转化为声音，让盲人可以“看见”东西。

瑞典天文学家卡尔·沙利耶（1862～1934）在富尼埃·达尔伯的书中投入了极大的热情，他将这个想法发展成了一个更加精细的牛顿宇宙学图景，避免了无限空间导致的矛盾。沙利耶采用了一种无穷分层星系群

的数学描述（图4.2），通过恰当安排使无限大宇宙的平均密度等于零！[5]这解决了长久以来关于夜晚的天空为什么总是黑暗的问题：在他的无限分层中，所有恒星的整体贡献可以忽略。



图4.1 卡尔·沙利耶

沙利耶发明的这种聚集图案后来被称作“分形”分布，这是分形的概念第一次被应用到了科学中。1972 年，贝努瓦·曼德尔布罗特（Benoît Mandelbrot，1924～2010）引入了分形的概念，用于描述一种不断地在更大的尺度上重复自身的图案。自然界里到处都能看到这种设计，分形可以让我们简单地理解大树的分枝和人体新陈代谢系统的图案（图4.3）。在任何情况下，自然选择都偏好聚集成分形的样式，因为其优点是能以最小的体积和重量来形成巨大的表面积（例如，用来吸收营养物质）。

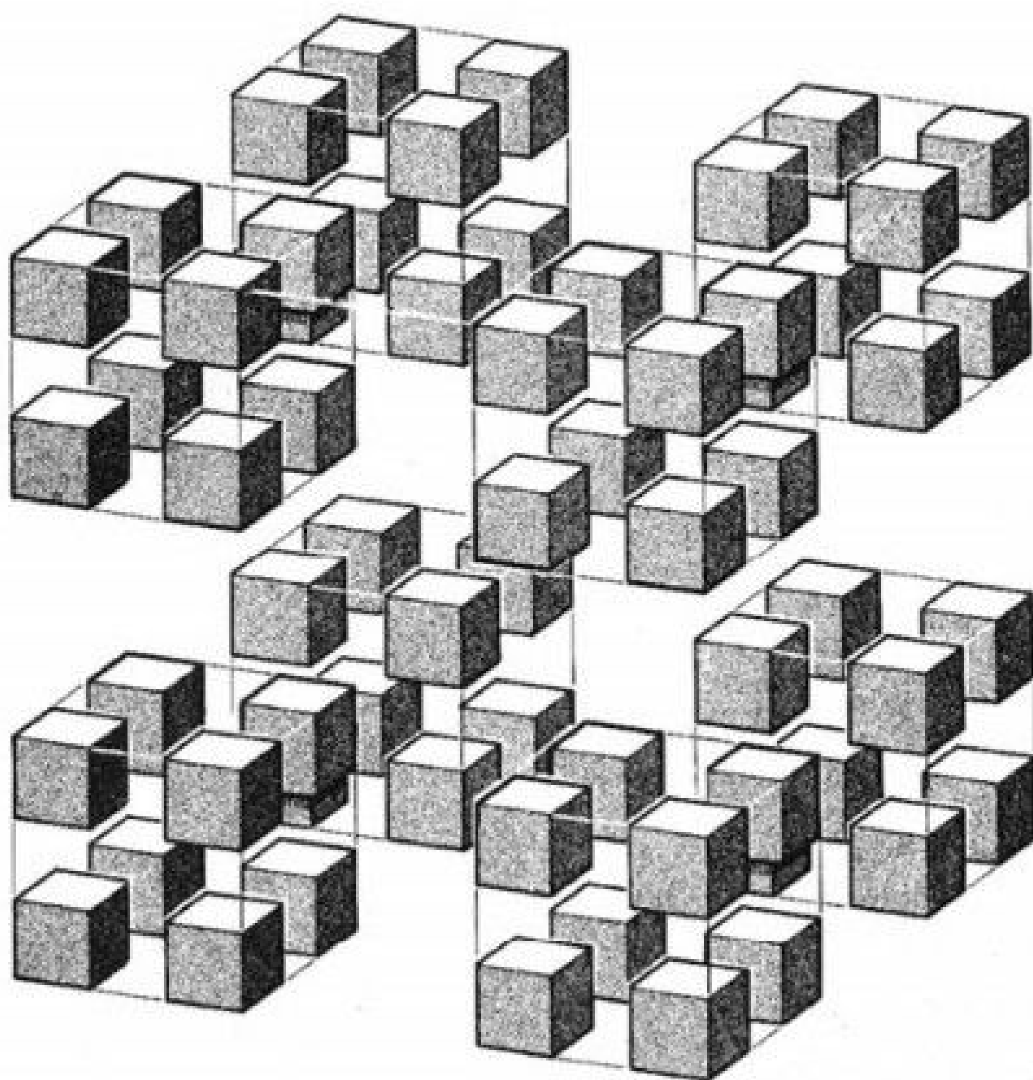


图4.2 这组分形分布显示了结团行为的三级结构。每一个方块都表示一个“星系”，每八个星系都聚集成一个大号的“星系团”，然后依次聚集成

一个“超星系团”。这个结团过程可以向上或向下无限类推下去

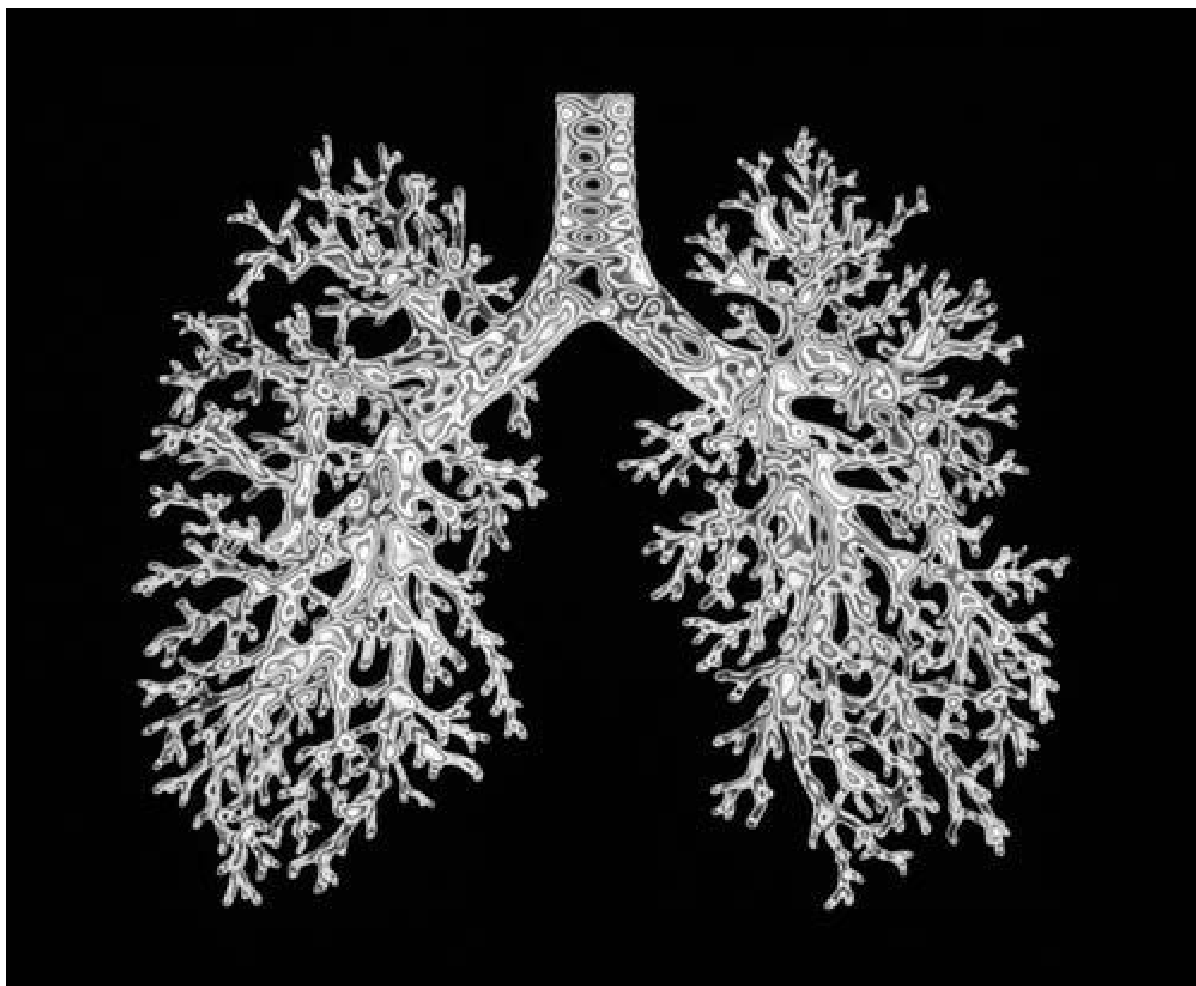


图4.3 人体肺部支气管系统的分形分布。气管不断地细分成更小的分支，或者叫支气管，这样就能在有限的空间中使得与空气的接触面积最大化

爱因斯坦用哥白尼原理简化了宇宙学模型，而沙利耶的宇宙是这个原理的真正扩充：在你看到的任何尺度上，它都表现出相同的图案。他仔细地建立了模型，因此尽管体积无穷大，其中聚集的物质也可以延伸到无穷远而不是只限于一隅。这种聚集减弱得足够快，于是所有恒星发光的总和不会让天空彻夜明亮[6]，宇宙每一处所受到的引力也是有限的，并且在所有尺度上，恒星受引力作用而产生的运动速度都很小。[7]

这种类型的宇宙学方案避开了牛顿理论在无限大宇宙中产生的问题，消除了爱因斯坦对用牛顿理论描述空间提出的异议。1922年，奥地利哲学家、自学成才的物理学家弗朗茨·塞利蒂（Franz Selety，1893～

1933?）[8]对这一模型给予了热烈的支持。就在那一年，塞利蒂将分层宇宙模型清晰地表示了出来，发表在了当时顶级的物理学期刊上。他的论文展示了如何解决爱因斯坦对无限大牛顿宇宙提出的所有异议。[9]简单地说，他开出了一剂无穷大分层宇宙的处方，恒星聚在一起总质量无穷大，填满了整个空间，平均密度却是零，也没有特殊的中心。[10]

但支持这种宇宙模型的人也面临着一个问题，那就是要解释物质一开始是如何按照这种"精细"嵌套分层的星系群结构排列的。爱因斯坦早已提出了这个忧虑。即使一开始这样的结构可以产生，也会很快分崩离析，形成其他的结构，因为当随机运动的恒星脱离原来的星系群时，它就会被临近星系群的引力俘获。塞利蒂不得不承认，一开始就产生这种结构的可能性微乎其微，但当你研究更高一层的星系群时，只要密度随着距离的增大而平方反比地减小，在将来有限的时间内，恒星的随机运动速度不足以使整个分层结构逐渐分离。[11]

爱因斯坦迅速回应了塞利蒂的论文[12]，说他没有解决任何问题，因为他错误地假设了牛顿宇宙学的必然性，并且承认"构建分层结构的宇宙是可能的……但无法令人满意"[13]。塞利蒂继续在1923～1924年间发表的论文中宣传他的宇宙模型[14]，并吸引了一些人的支持，包括法国杰出的数学家埃米尔·波莱尔（Émile Borel），但爱因斯坦后来对这个模型再也没有作出评论。

分层宇宙只风靡了很短的时间，尽管事后看来塞利蒂的计算完全正确。20世纪70年代和90年代，这种模型偶尔会在讨论爱因斯坦理论的研究中出现几次，用来解释当时观测到的星系群结构数据。[15]如今，对微波背景辐射的温度涨落的精确观测告诉我们，星系群不会形成无限层的分形结构。[16]

卡斯纳博士的宇宙

希望我永远不会犯下错都错不到点上的那种错误。

——萨米尔·萨玛吉（Samir Samaje）

也许你常常听说一种叫做万维网的东西，而且毫无疑问，一种名叫Google的搜索"引擎"更是如雷贯耳，它把触手伸向了信息世界的每一个角落，能够以令人惊叹的速度迅速搜出一套餐具或一本书。对于一个开展计算机密集型、海量搜索业务的公司来说，Google是一个奇特的名字，而它的总部名称Googleplex，听起来更是古怪。

这些名字背后的故事来自于一位美国数学家爱德华·卡斯纳（Edward Kasner，1878～1955），纽约哥伦比亚大学伯纳德学院的教授。除了在数学众多领域内的研究工作，他还热心于通过演讲、写书和撰文向公众和年轻人介绍数学知识。他最有名的一本书叫《数学与想象力》（*Mathematics and the Imagination*），是和詹姆斯·纽曼（James Newman）合写的，于1940年首印，直到现在还在重印。其中一章谈到

了大数，并给出了一个简洁而又巨大数目的例子：10^{100}，也就是1后面跟了100个0。（相比之下，整个可见宇宙中只有大约10^{80}个原子和大约10^{90}个光子。）



图4.4 爱德华·卡斯纳

1938年，卡斯纳 9岁的侄子弥尔顿·瑟罗塔（Milton Sirotta），给这个数起名叫Googol，并创造出一个无法想象的大数形式Googolplex，也就是把Googol放在10的指数位置上，于是：

$$1 \text{ Googolplex} = 10^{\text{Googol}}$$

这个数太大了，如果我开始写出它的全部数字，10000000000……可观测的宇宙只有10^{29}厘米长，根本装不下。

根据计算机科学家戴维·科勒（David Koller）所述[17]，在 1996年，当时斯坦福大学计算机系的两个博士生拉里·佩奇（Larry Page）和谢尔盖·布林（Sergey Brin），正开始研究如何描绘互联网页面中普通词语和链接之间形成的关联网络。最终，他们研究出的网页排名算法变成了世界上最有效的互联网搜索引擎。起初他们管这个新技术叫“搓背”（BackRub），但后来几年他们想换一个更好的、能体现搜索结果中海量数目链接的名字。他们的一个同学肖恩·安德森（ Sean Anderson），突然提出可以叫Googolplex，这被佩奇简化成了Googol。这名字听起来不错，所以安德森立刻在电脑上搜索，看Googol.com的域名有没有被人注册过。他急匆匆地一搜，把域名误拼成了Google.com，发现还能注册。布林好像挺喜欢这个新的（错误的）写法，觉得比以前的好，于是在 1997年 9月 15日那天，以他和佩奇的名义注册了Google.com。后来，当谷歌成长为一个大公司时，其位于加州圣何塞市附近圣克拉拉郡山景城的别具一格的总部大楼也被冠以了Googleplex的昵称。

除了向世人宣传一些大数的新名称之外，爱德华·卡斯纳还做过别的事。1921年，他是寻找爱因斯坦方程组新的解的一小群人中的一员。他了解爱因斯坦和德希特早先的那些解，它们都考虑了起排斥作用的宇宙学常数。虽然不像那些著名的先驱者，卡斯纳对天文学不甚了解，但他精通爱因斯坦方程组背后的抽象数学，并打算从纯数学的角度去挑战它，寻找新的忽略了物质影响的宇宙解。这是德希特采用过的假设，但和德希特不同，卡斯纳从方程组中完全剔除了宇宙学常数项。为了弥补这些简化的缺陷，卡斯纳引入了一个全新的假设：宇宙的不同方向能够以不同速率膨胀。卡斯纳的各向异性宇宙的空间几何是平坦的、欧几里得的。它的体积有限，从过去某个有限的时刻开始，并将永远膨胀下去。弗里德曼和勒梅特的宇宙某些部分的膨胀是球对称的，相比之下，卡斯纳的宇宙每一个方向都以不同的速率膨胀，像一个椭球面。

卡斯纳的宇宙有一个令人震惊的特点。尽管它的总体积在增加，但它实际上只在某两个维度的方向上是膨胀的，而在另一个垂直维度的方向上则是收缩的（图 4.5）。因此，在卡斯纳的宇宙中，一个空心球面的赤道地区会越来越大，越来越像一个椭球，而同时两极地区在朝着中心收缩。宇宙的形状会越来越像一块薄饼。[18]

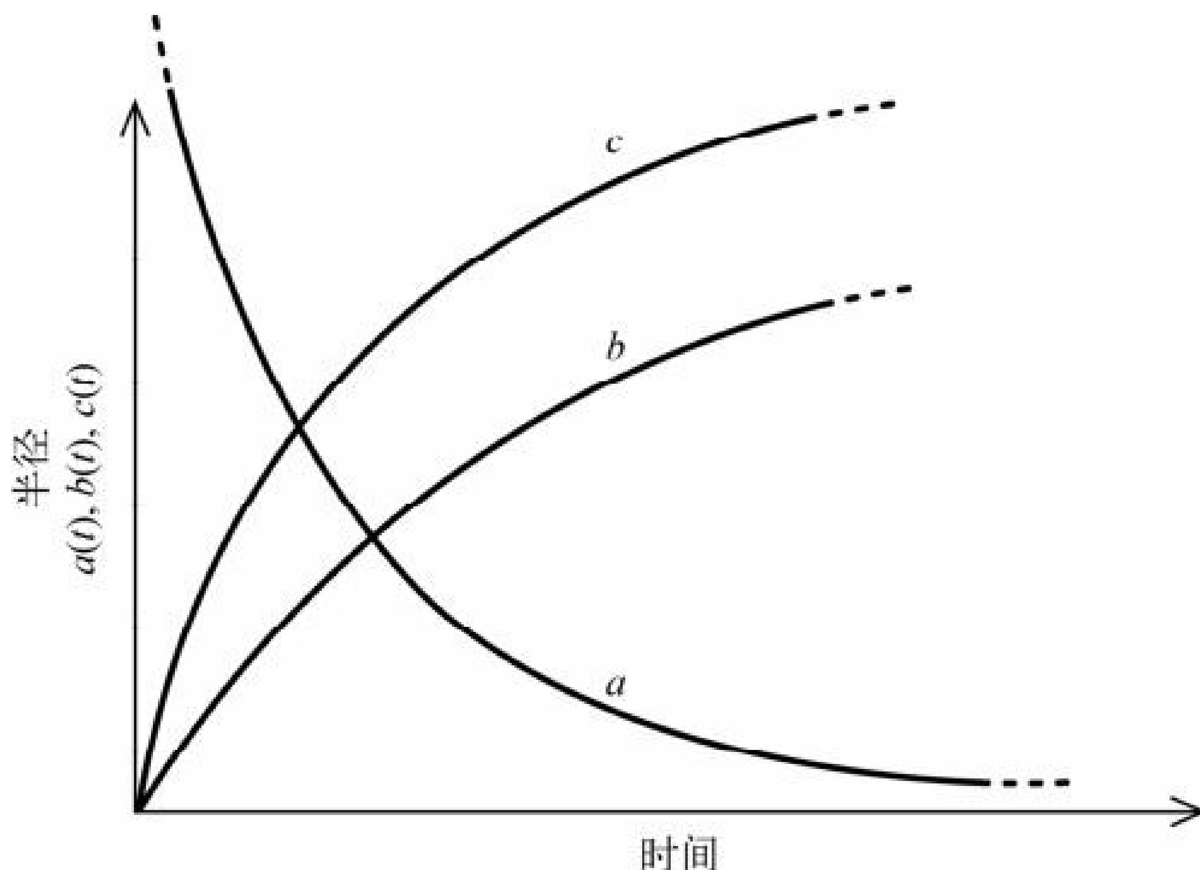


图4.5 卡斯纳宇宙三个维度X、Y和Z的膨胀。其中两个膨胀时，第三个就收缩，总体积XYZ总是正比于时间

这是一个奇怪的宇宙。它不包含物质，空间并没有弯曲，但又在膨胀。这是一个由不同方向之间的膨胀速度差异所驱动的世界。我们很熟悉引力差异导致的效应，因为从潮汐中我们就能体会到这一点。月亮（和太阳）对地球上靠近它们一面的海水产生的拉力更大，所以这一侧的海平面比另一侧更高。地球自转导致潮汐的强度规律地变化，并且遵循立方反比而不是牛顿的平方反比定律。卡斯纳的宇宙存在这种潮汐力，因此它是各向异性的。它允许一个宇宙的不同方向“启动”不同速率的膨胀。[19]

卡斯纳的宇宙看起来非常特别。如果往其中加入一些物质，它就会

逐渐向爱因斯坦和德希特的各向同性宇宙演化。如果往其中加入起排斥作用的宇宙学常数，不管其中有没有物质，它都会向德希特的指数膨胀宇宙演化。[20]然而，如果我们逆着时间回溯到它的“开端”，t＝0的时刻，那时物质、辐射和宇宙学常数都不会对它的面貌产生任何影响。膨胀速度的差异最终胜出，宇宙开始膨胀，它开始时是伸长的“针”的形状，体积为零，不是一个“点”，而是在一个维度的尺寸无穷大，另外两个与之垂直的维度的尺寸为零。尽管它的性质如此特殊，但事实证明，卡斯纳的宇宙会对理解可能存在的宇宙产生巨大的影响。

狄拉克的宇宙——引力在其中持续衰败

我觉得你不可能在研究物理学的同时又写诗。做科研时，你想要说一些别人都不知道的东西，又要让所有人都能理解。而创作诗歌时，你必须说一些别人都已经知道的事，又要说得大家都看不懂才行。

——保罗·狄拉克[21]

保罗·狄拉克（Paul Dirac，1902～1984）在剑桥大学担任卢卡斯数学教授时，正好与爱丁顿在剑桥大学天文台生活和工作的时间有所重合。狄拉克可以被称作是20世纪英国最伟大的物理学家，他建立了量子力学中的许多理论，规范了量子力学的数学表述形式，预言了反物质，发现了特定基本粒子的统计性质以及重要的、描述相对论性电子行为的“狄拉克方程”。1933年，他年仅31岁，就成为诺贝尔物理学奖的最年轻得主①，而一年前他就已经被授予卢卡斯教授教职。

① 最年轻的得主应该是威廉·劳伦斯·布拉格（William Lawrence Bragg，1890～1971），1915年获奖时只有25岁。——译者注

狄拉克的传记作者[22]把他描述为“最奇怪的人”，而与他同在剑桥的那些人几乎不敢说自己“了解”他。他沉默寡言，寥寥数语就能将别人的特点概括出来。例如，路德维希·维特根斯坦就被狄拉克说成是“可怕的人，喋喋不休”。[23]关于狄拉克生活朴素、井井有条而社交能力低下的故事非常多，因此1937年2月，他刚度完蜜月就向《自然》杂志投稿，出乎意料地进军宇宙学领域，就显得一点儿也不奇怪了。10个月以后，一篇更长的文章出炉，提出了宇宙学的一个新原则。之后他再也没有写过这方面的文章，直到 35年后又重新捡起了这个题目，仿佛中间的 35年不存在似的。

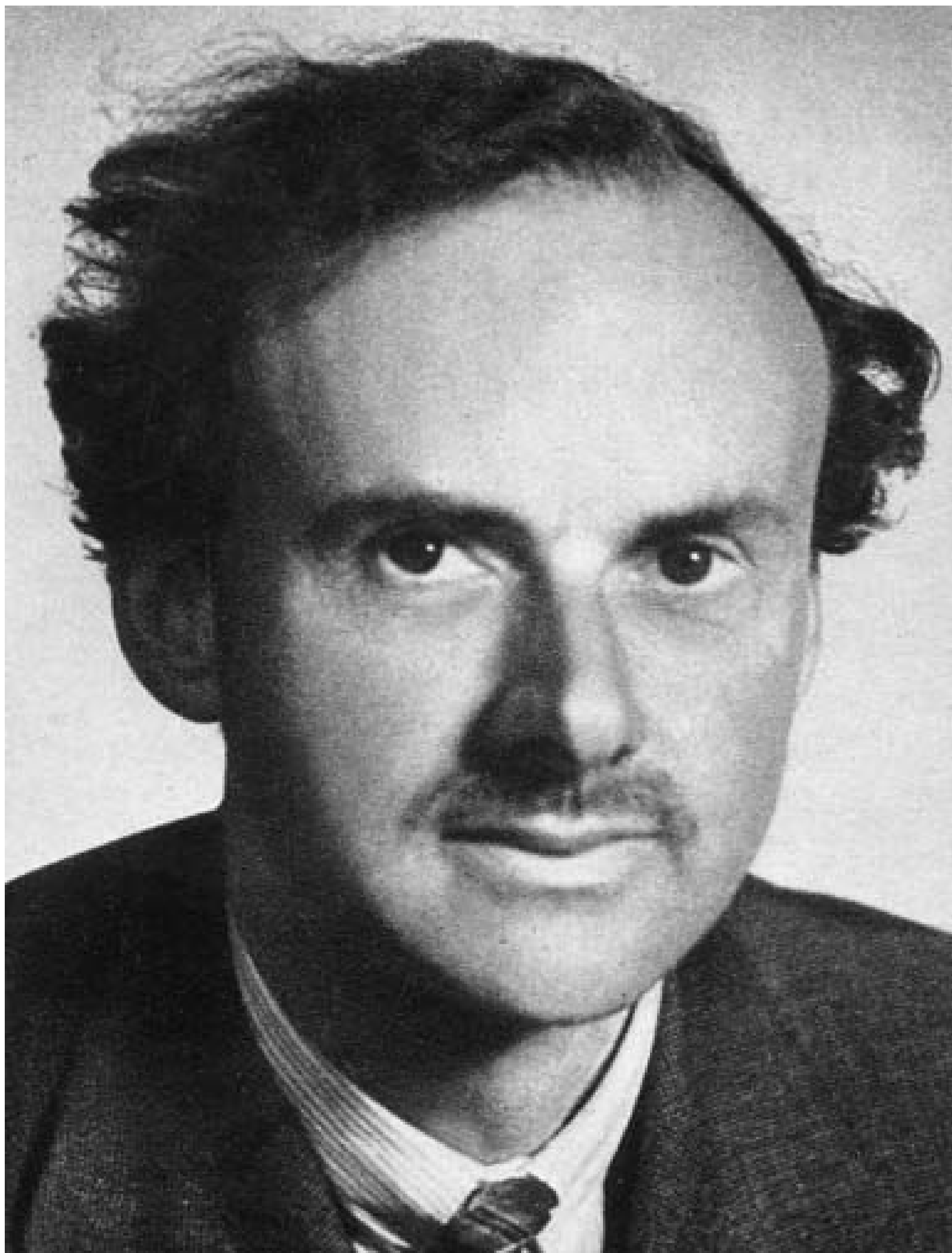


图4.6 保罗·狄拉克

当时大部分的物理学家，一听说狄拉克写了一篇宇宙学的研究论文，都非常期待那是一个全面的新理论，或是一种用高超的数学技巧求解爱因斯坦极富挑战的方程组的新方法。恰恰相反，狄拉克的想法虽然非常令人着迷，但又非常简单，很多人感觉是有点古怪，甚至他的朋友尼尔斯·玻尔（1885～1962）打趣说，“看看人结了婚就变成什么样了”。狄拉克认为，如果我们在物理学中碰到巨大的无量纲量，例如10^{40}或10^{80}，它们不太可能是相互独立、无关的，很有可能存在一个未被揭示的自然数学法则，会将这些量联系在一起。这就是狄拉克的大数假设。[24]狄拉克发现的大数共有三组，其中包括了宇宙的年龄 t、光速 c、电子的质量m_e、质子的质量m_p以及牛顿引力常数G。从这些量中他构造了以下三个数：

$$N_1=\text{可观测宇宙的尺寸和电子半径的比值}=ct/(e^2/m_ec^2)\approx 10^{40}$$

$$N_2=\text{一个电子和一个质子之间电磁力和引力的比值}=e^2/\mathrm{G}m_em_p\approx 10^{40}$$

$$N=\text{可观测宇宙中质子的总数}=c^3t/\mathrm{G}m_p\approx 10^{80}$$

根据他的假设[25]，N_1、N_2和$\sqrt{N}$在非常精确的近似下很可能是相等的。如果这些大数没有任何联系的话，才是非常非常奇怪的事。狄拉克相信，自然界中肯定存在未知的法则，使得 $N_1\approx N_2$或$N\approx N_2^2$这样的公式（或近似的公式）成立。

狄拉克的大数假设抛出的激进意涵是，它要我们相信，一系列自然界的经典常数必然随宇宙年龄t的增加而变化，因为他要求

$$N_1\approx N_2\approx\sqrt{N}\approx t$$

所以，这三个经典自然常数的组合并不是常数，而必然随着宇宙年龄t的增加而缓慢地改变，于是：

$$e^2/\mathrm{G}m_p\propto t \quad (*)$$

① 符号∝表示左边的表达式与右边呈正比。——译者注

狄拉克抛弃了牛顿宇宙学常数G的不变性，而选择坚持自己的假定。他提出，在宇宙的时间尺度上，牛顿引力常数与宇宙年龄呈反比，

如并且满足方程（*）的要求。于是，相比现在测量到G的数值，从前的G会更大，而将来的G会变小。你会发现$N_1 \propto N_2 \propto \sqrt{N} \propto t$，也就是说，它们三个之所以数值巨大是因为宇宙年龄很古老[26]：随着时间的流逝，这三个数越来越大。[27]

$$G \propto 1/t$$

狄拉克的结论有三个关键要素。首先，他试图证明，以前人们认为是巧合的事情实际上是一些被忽略的深刻联系的结果。其次，他断定，空间曲率和宇宙学常数必然是零，否则这些数值会不断变大。最后，他牺牲了G的不变性，这是已知的最古老的自然常数。不幸的是，这些假设存活的时间并不长。如果G的数值都能发生变化，那就坏事了。狄拉克假设，从前引力相互作用的强度很强。这会导致太阳的输出功率发生变化，而从前的地球会比通常认为的热得多。[28] 1948年，美国物理学家爱德华·泰勒（1908～2003）说，若是那样的话，寒武纪之前的海洋会沸腾，我们现在所知的生命也就不可能演化至今。[29]

泰勒的朋友乔治·伽莫夫回应了沸腾海洋的问题。他提出，如果假设狄拉克的巧合是因为电子电荷 e在随时间变化，情况就会改善许多，方程（*）要求e^2的数值随着时间增大。[30]

不幸的是，这个假设存活的时间也不长。如果伽莫夫变化的e的想法成立，那么将由此导致种种使生命在地球上无法出现的糟糕后果。没过多久就有人发现，他的理论会导致太阳在很久以前就耗尽了所有的核燃料。如果e^2的数值与宇宙的年龄呈正比的话，现在的太阳早就不可能发光了。

除去这些失败的尝试不说，狄拉克的想法还是启发了我们，去探索宇宙学模型的新可能。改变引力常数，让它随时间变化，相当于说爱因斯坦的广义相对论不正确，或者不完备。你不能像狄拉克那样，若无其事地把一个常数变成一个变量。在爱因斯坦的引力理论中，任何形式的能量都产生引力。这些能量弯曲了空间，决定了时间流逝的速率。改变G 的大小也会产生相同的效果。后来，一些物理学家试图使整个想法建基于一个坚实的基础之上，用爱因斯坦理论的小规模修正版来解释引力“常数”变化的原因，把它看作一种能量和引力作用的新来源。

1939年，第一个用这种想法构建了完整理论的是个不知名的德国物理学家，帕斯库尔·约尔丹（Pascual Jordan，1902～1980）。他拓展了爱因斯坦的相对论，向其中加入了一个新的能量场，表示传统意义上的引力常数G的变化。约尔丹和马克斯·玻恩（1882～1970）、沃纳·海森

堡（1901～1976）一起发表了一系列论文，对量子力学的发展作出了重要贡献。但很不幸，约尔丹是一个狂热的纳粹分子，在1933年希特勒掌权的那一年就加入了纳粹党，继而在 1934年加入了纳粹党冲锋队。[31] 1939年，他加入了空军，并在第二次世界大战中开展气象学的研究工作，还在靠近波罗的海的皮纳穆德 V-1和 V-2火箭研究机构待了一段时间。但不管他在武器系统的研究中表现得如何狂热，他的上级看起来并不信任他，可能是因为他以前与犹太物理学家有交往，如玻恩。约尔丹的政治倾向使得他与其他物理学家之间产生了一个巨大的鸿沟。直到 20 世纪50年代早期，在他恢复了一个学术职位时，这事才算平息。因此，在当时，约尔丹前面十几年的宇宙学研究都被大大忽略了。有人说，他的政治活动使他丧失了分享1954年诺贝尔物理学奖的机会。

尽管狄拉克对宇宙学的跨界研究很短暂，并随后被更精确的如约尔丹的表述所取代，但狄拉克发现的数字巧合后来有了一个有趣的解释。罗伯特·迪克（Robert Dicke，1916～1997）指出，$N \approx N_1^2$的公式就是在说，我们观测到宇宙的时间t（我们叫做“年龄”），大体上与一颗恒星开始形成、步入稳定演化的阶段（在核聚变反应中燃烧氢元素产生氦元素）的时间相当。既然我们不可能在恒星形成以前就存在，也不能在恒星熄灭以后还活着，这种巧合就根本没什么好大惊小怪的。[32]我们不太可能活在一个违背“巧合”（*）的宇宙。

狄拉克并没有意识到，我们在观察宇宙时，正处于宇宙历史的一个特殊时期。对于一个膨胀的宇宙来说，只有在一些特定的阶段中，生命才有可能存在，而我们也只有在宇宙历史的那个宜居时期才能进行天文学研究。

狄拉克回应了迪克的重要反对意见，承认在一个膨胀宇宙之前的某个时期，生命是不可能存在的，因为宇宙太热太致密了；但是他相信，一旦出现，生命将在宇宙中永远延续下去。[33]在狄拉克看来，生命并没有被局限在宇宙的一段历史时期中。他相信，它的未来无限。

爱因斯坦和罗森的波光潋滟的宇宙

……空间的一小部分实际上可以类比于一个大体看来平坦的表面上的突起，也就是说，通常的几何规律在这里失效了。这种弯曲或扭曲的性质以波的形式从空间的一部分传播到另一个部分，绵延不绝。

——威廉·金登·克利福德[34]

1932年，爱因斯坦离开欧洲，去了新泽西州新建成的普林斯顿高等研究院。路上他在牛津大学做了短暂的停留，并在基督教堂学院做了个研究报告。远离了危机重重的中欧，他重新开始考虑求解他的方程组。1935年，他得到了一个年轻的研究助理纳森·罗森（Nathan Rosen，1909～1995）的帮助，帮他进行数学计算。在随后的两年里，罗森和爱因斯

坦合作撰写了许多理论物理学领域中的著名论文。[35] 1936年，爱因斯坦和罗森发现了爱因斯坦方程组的一类新的解。它描述了一个膨胀的圆柱状宇宙，因此所有的事情在随时间变化的同时，也沿着空间中的某一个方向变化。这种形状简化了纷繁复杂的爱因斯坦方程组，使人们得以找到一个精确解。这种宇宙有一个令人惊诧的新特点，这是以前从爱因斯坦方程组找到的其他宇宙模型中从来没有过的。这种宇宙的空间中传播着种波，所到之处的空间几何纷纷掀起涟漪。这有点像给卡斯纳的宇宙（在不同方向上以不同的速率膨胀）增添了一些波，它们从对称轴向外传播，就像一筒厨房卷纸松开的时候向外一层层地脱去。

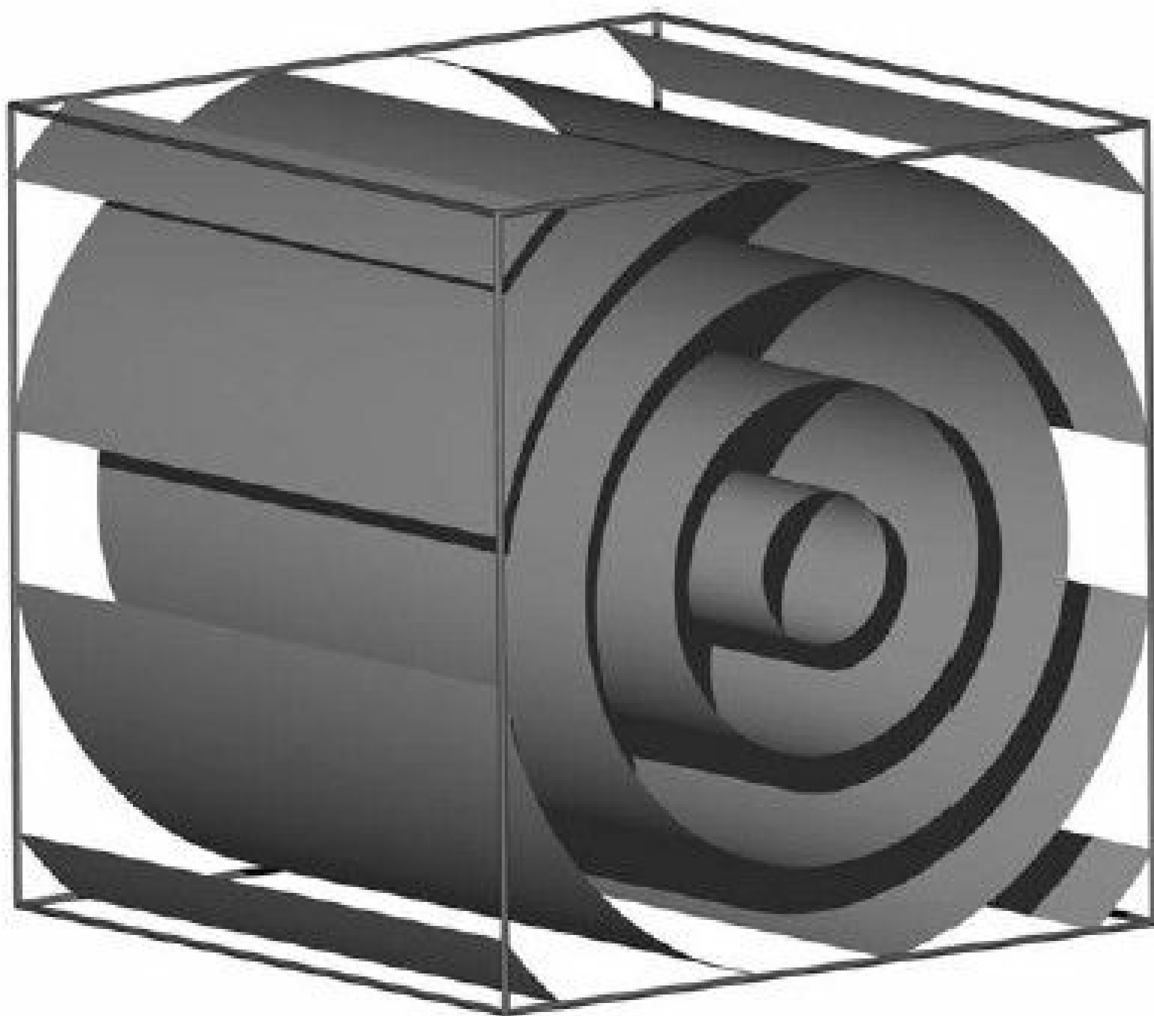


图4.7 爱因斯坦–罗森宇宙包含了圆柱形的引力波，引力波从膨胀宇宙的一条线中发出，向外传播

对爱因斯坦和他同时期的人来说，这种宇宙的最有趣之处是存

在“引力波”。爱因斯坦–罗森宇宙不包含物质，因此任何波动都会掀起空间几何的涟漪，随时间传播开去（图 4.7）。这种引力波的想法曾经风行了一段时间，也引起了一些争议。有些人相信这只不过是在“纸上谈波”，它们并不对应空间中任何真实存在的起伏，只不过是爱因斯坦方程组在特殊的坐标系下产生的结果。有些人则认为这是真实存在的波，如果一束波朝你袭来，它会对你产生影响（沿着一个方向拉伸你，并沿着垂直的方向挤压你，就像潮汐力）。

很快，爱因斯坦和罗森就意识到，这个解是一个绝佳的理论检验对象，要想平息这场争论并不需要借助任何不确定的近似，也不需要数值计算。值得注意的是，他们首先得到的结论是，这种柱状波并非“真实存在”，它只不过是他们选择的坐标系的人造产物。他们说，这就像是一个地理学家的地球仪，所有的经线都在两个极点处相交。对于不懂的人来说，这可能说明地球上的那个地方大事不妙，事物都汇聚在了一起，但事实上什么事都没发生。在地图的极点处，你总是可以换成别的坐标系，事物就变平常了。想象一下，你把经线表示成歪歪扭扭的线。于是，对于不懂的人来说，这些线意味着地球上有一些东西在不断波动——肯定又是你错了。然而，要小心了，如果你看的是一幅等高线图，你会发现许多复杂的波浪线看起来非常类似。在这种情况下，如果你断定这些线并不对应地球上任何“真实存在”的东西的话，你就错了。爱因斯坦和罗森当时面临的问题是，他们应该把这些曲率波归为像等高线一样的“真实存在”，还是像歪歪扭扭的经线一样纯属人造产物？

1936年夏初，他们向美国的顶级物理学期刊《物理评论》寄出了自己的论文。文中声称，他们的新宇宙学模型中的引力波并非真实存在。6月 1日，他们的文章被签收了，之后就像其他的投稿一样，被编辑寄给了另一个科学家进行评估。7月23日，论文的审稿意见回来了，并寄给了爱因斯坦，按照惯例其中没有透露审稿人的姓名。现在我们知道给编辑提供审稿意见的不是别人，正是霍华德·罗伯特逊（Howard Robertson， 1903～1961），一位对广义相对论技术细节了如指掌的美国科学家。他不相信文章的结论，并一针见血地指出了作者在文中得出错误结论的地方。他明确地认为，这个宇宙模型中的引力波是真实存在的，并要求作者考虑他的这个意见。在说到爱因斯坦的回应之前，请注意爱因斯坦以前所有的工作都发表在欧洲期刊上，当时欧洲并没有像这样对论文进行同行评议的传统。欧洲期刊要么看作者的名气大不大，要么看有没有功成名就的科学家做介绍人，要么编辑亲自审稿。拒绝一个功成名就的作者投稿相当不礼貌，这种事极少发生。结果，由于不理解美国的审稿机制，爱因斯坦在听说编辑把他的文章寄给了另一个科学家的事以后很沮丧。他给《物理评论》的编辑约翰·塔特（John Tate）回

信说：

我们（罗森先生和我）已经向您寄去了我们的稿件，用于发表，并没有授权您在付印之前可以给别的专家审阅。我认为没有必要回应您那位匿名专家的意见（本来就是错误的）。鉴于此，我宁愿选择在别的地方发表这篇文章。[36]

很快，爱因斯坦将这篇文章投到了《富兰克林研究所杂志》，以前他在那儿发表过一篇论文。

过了几个月，爱因斯坦仍然相信自己的新宇宙模型不包含真实的引力波。后来，一切都改变了，因为爱因斯坦与罗伯特逊交上了朋友——就是在《物理评论》匿名给爱因斯坦审稿，并对文章结论提出质疑的那个罗伯特逊。罗伯特逊说服了爱因斯坦，相比他和罗森得到的结果，有一种解的形式更清楚，毫无疑问，其中存在的柱状引力波是真实的。那个时候，爱因斯坦和罗森的论文已经被《富兰克林研究所杂志》接受了，不过爱因斯坦在收到校样稿时，得以有机会在结论中做一个保全面子的改动，增加了一段备注，感谢“我的同事罗伯特逊教授在澄清原文中的错误时给予的亲切帮助”。[37]

他们刚投了稿，纳森·罗森就已经前往苏联的基辅大学工作了。于是，他是在报纸上读到爱因斯坦发表了新论文这个有新闻价值的事件时，才听说了后面发生的事。罗森并没有被爱因斯坦和罗伯特逊的观点说服，不相信引力波的真实性，随后就独自发表了一篇文章，仍然坚持以前那个错误结论。罗森一直和他们争论，直到20世纪70年代，那时几乎所有的物理学家都早已经相信引力波的真实性了。1957年，爱因斯坦去世两年后，在北卡罗来纳大学教堂山分校的会议上，理查德·费曼（Richard Feynman，1918～1988）提出的一个简单观点成了对罗森的关键反驳。费曼证明，如果引力波垂直经过一根粗糙的棍子，附着在棍子表面的水珠（由于存在摩擦力，它们被称作“黏性水珠”）会前后移动。水珠在棍子的粗糙表面移动会导致摩擦生热，就像你取暖时摩擦自己的双手一样。温度升高说明引力波就是热源，因而必然携带能量。所以，引力波并不只是纸上谈兵。

费曼的“黏性水珠”之说化解了所有对引力波真实性的质疑。引力波就像潮汐力。如果一束引力波垂直穿过这一页书，就会沿着一个方向拉伸它，而沿着垂直的方向挤压它：正方形会变成长方形，而圆形会变成椭圆形。

关于这个故事还有件有趣的轶事。费曼好像看不上教堂山会议的议程，注册时使用了假名，于是“黏性水珠”之说是匿名发表的。在他的回忆录《别闹了，费曼先生》之中，他讲述了寻找会议地点的经过①：

① 此处主要是在吴程远的《别闹了，费曼先生》译本上进行少量修改。——译者注

1957年间，我去北卡罗莱纳大学参加一个讨论引力的研讨会。我本来要以物理学其他领域的专家身份来讨论引力理论。（由于我无法参加第一天的会议）当我的航班在机场降落时，已经是会议的第二天了。我走到叫出租车的地方，跟那人说："我要到北卡大学。"

"你说的是哪一所，"他说，"在罗利的北卡州立大学呢，还是在教堂山的北卡大学？"

不用说，我完全搞不清楚。"它们在哪里？"我问，心想两个地方应该挨得比较近。

"一家在北方，另一家朝南走，路程都差不多远。"

我身上没带任何资料能让我弄清楚究竟是哪个地方，而旁边也没有像我那样晚了一天才来开会的人。

我灵机一动。"听着，"我跟出租车站的人说，"会议是昨天开始的，所以昨天一定有很多人路过这里去参加研讨会。让我形容一下这些人，看你有没有印象：他们多半有点迷迷糊糊的，边走边谈，不大理会自己究竟往哪个方向走，谈话内容都是'姬—谬—拗，姬—谬—拗'的。"

他整张脸都亮了起来，说："你要去的是教堂山！"他挥手招来出租车，"带这位先生去北卡教堂山分校。"

"谢谢！"我说，然后顺利地抵达了会议地点。[38]

注释

[1] G. K. Chesterton,*Orthodoxy*, London (1908).

[2] 设想无穷多的物质平均分布在无穷大的空间中。你可以假设自己受到了引力，大小和方向随便选。根据你的选择，你可以划出一块球形空间，球心的方向和引力的方向相同，球的半径使得其中包含的物质所产生的引力刚好满足你的要求，而你就位于球体的表面。此时，牛顿会告诉你，你只会感受到球体内物质产生的引力，外面的物质产生的总引力为零。所以，你所受到的引力正是你开始时假设的引力。用这种方法，你可以"证明"你想受到多少引力，就受到了多少引力。这叫做引力佯谬（Gravitational Paradox）。

[3] E. E. Fournier d'Albe, *Two New Worlds: The Infra World. The Supra World*, Longmans, Green & Co., London (1907).他在1906年出版的《电子理论》（*The Electron Theory*）一书中已经对这个想法做了更详细的阐述。

[4] E. A. Poe,*Eureka – A Prose Poem, Putnam*, New York (1848).关于爱伦·坡对"其他宇宙"的描述，进一步的讨论参见：E. R. Van Slooten,Nature 323, 198 (1986); A. Cappi, *Quart. J. Roy. Astron. Soc.* 35, 177 (1994).

[5] C. Charlier, 'How an Infinite World May Be Built Up',*Arkiv f. Matematik och Fysik* 16, no. 22 (1922).

[6] 这个问题被叫作奥伯斯佯谬，是由英国天文学家埃德蒙·哈雷（Edmund Haley，1656～1742）首先注意到的。

[7] 对此有一篇综述写得不错，参见：J. D. Norton,*The Expanding Worlds of General Relativity*, Birkhäuser, Boston (1999), pp. 306–308.

[8] 塞利蒂的个人资料很难找，因为他由于某种原因改过姓，1918年以前叫弗朗茨·约瑟夫·雅伊特勒斯（Franz Josef Jeiteles），参见：T.

Jung,*Acta Historica Astronomiae* 27, 125 (2005)。注意，塞利蒂（Selety）其实只是把雅伊特勒斯（Jeiteles）倒过来写，这大概是怕被原来姓中“-eles”的后缀暴露了其犹太人的身份。当时在奥地利，反犹主义甚嚣尘上。

[9] F. Selety,*Annalen der Physik* 68, 281 (1922).

[10]只有当空间的体积趋于无穷大时，平均密度才趋近于零。当你用总质量除以总体积时，平均密度会随着体积的增大而迅速减小，比你能想到的任何不为零的数都小。

[11]随机的速度v和结构的尺寸r、密度ρ(r)之间的关系为 $v^2 \approx GM/r \propto G\rho(r)r^2$。因此， 如果物质的密度ρ(r)随着尺寸的增大而减小，如 $\rho(r) \propto r^{-2}$或者衰减得更快的话，随机的速度v并不会随着r的增加而增加。塞利蒂倾向于选择 $\rho(r) \propto r^{-2}$的特殊情况。

[12] A. Einstein, *Annalen der Physik* 69, 436 (1922).塞利蒂对此的回应，参见：*Annalen der Physik* 72, 58 (1923).

[13]可惜，爱因斯坦似乎被一种陈旧的观念影响了，这种观念认为没有证据表明宇宙中存在河外星系。

[14]最后一篇是：F. Selety, *Annalen der Physik* 73, 291 (1924).

[15] G. de Vaucouleurs,*Science* 167, 1203 (2000); J. R. Wertz,*Astrophys. J.* 164, 229 (1971); W. Bonnor,*Mon. Not. R. Astron. Soc.* 159, 261 (1972); C. Dyer,*Mon. Not. R. Astron. Soc.* 189, 189 (1979).

[16]有些人持不同意见，例如：F. S. Labini, ‘Characterising the Large-Scale Inhomogeneity of the Galaxy Distribution’, http://arxiv.org/abs/0910.3833.

[17]参见他的个人主页：http://graphics.stanford.edu/~dk/google_name_origin.html.

[18] x、y、z三个相互垂直的维度随时间的变化率是：

$$R_x = t^p \ , \quad R_y = t^q \ , \quad R_z = t^r$$

其中爱因斯坦方程组要求p、q、r三个数满足

$$p + q + r = 1 \text{ 以及 } p^2 + q^2 + r^2 = 1$$

这说明p、q、r的取值范围被限制在如下的不等式中：

$$-1/3 \leqslant p \leqslant 0 \leqslant 1/3 \leqslant q \leqslant 2/3 \leqslant r \leqslant 1$$

宇宙的体积$R_xR_yR_z = t^{p+q+r} = t$，因此是随时间增大的。但是 p小于 0，所以卡斯纳宇宙的y方向和z方向膨胀时，x方向一定在收缩。

[19]如果你要求在卡斯纳的宇宙中，每个维度的膨胀速率都相同，那么 p、q、r之间就相等了，就违反了上述公式。实际上，我们可以要求p=q=0，r=1。结果就成了狭义相对论里的平坦时空，即闵可夫斯基时空，只不过不是通常的坐标系。这个解是可以存在的，但不能描述膨胀的宇宙。

[20]这种宇宙模型是由欧托·海克曼（Otto Hechman）和英格丽伯特·舒金（Englebert Schücking）在 1959年发现的（*Handbuch der Physik*, vol. 53, Springer, Heidelberg, 1959, pp. 489–519），即

$$R_x = t^p (t+T)^{2/3-p}\ ,\quad R_y = t^q (t+T)^{2/3-q}\ ,\quad R_z = t^r (t+T)^{2/3-r}$$

其中T是一个常数，可以任意取值。简单地说，T表示卡斯纳的各向异性膨胀开始向各向同性转化的时间。我们发现，当t 远小于T时，这个宇宙就像卡斯纳的宇宙，但是当

t远大于T时，又变得越来越像爱因斯坦–德希特宇宙，即$R_x=R_y=R_z=t^{2/3}$。当我们向其中加入宇宙学常数Λ后，结果就是$R_x = [\sinh(At)]^p \times [\sinh(At) + At\cosh(At)]^{2/3-p}$，其中 T是常数，$A = (3\Lambda)^{1/2}/2$也是常数。关于$R_y$和$R_z$的公式可以将p分别替换成q和r后得到。当t接近0时，这种宇宙又回到了卡斯纳的宇宙；当t很大时，这种宇宙就趋近于德希特的宇宙，即$R_x = R_y = R_z = \exp\left[t(\Lambda/3)^{1/2}\right]$。

[21]狄拉克对 J.罗伯特· 奥本海默做出的评论，转引自：G. Farmelo, The Strangest Man, Faber, London (2009), p. 121.

[22] G. Farmelo,*The Strangest Man*, Faber, London (2009).

[23] G. Farmelo,*The Strangest Man*, Faber, London (2009), p. 220.

[24] P. A. M. Dirac,*Nature* 139, 323 (1937) and *Proc. Roy. Soc. A* 165, 199 (1938).“自然界中的任何两个巨大的无量纲量之间的数学关系必然很简单，其系数在相同的数量级上。”爱丁顿曾经想要解释物理常数为什么拥有特定的取值，好像因而使狄拉克注意到了某些与物理定律和宇宙结构息息相关的神秘的数。

[25]大数相等的假设本来并不是狄拉克最先提出的。爱丁顿和别的人以前曾经写过这样的关系式，但是爱丁顿并没有区分整个宇宙中的粒子数（可能是无限多的）和可观测宇宙中的粒子数，后者是一个半径等于宇宙的年龄乘以光速的球形空间中的粒子数。

[26]$N \propto t^2$的结果让狄拉克错误地认为，这说明宇宙在不断地创造新

的质子，参见：P. A. M.Dirac,*Proc. Roy. Soc. A* 333, 403 (1973).但实际上，这只不过说明随着宇宙年龄的增加，可见宇宙的视界也在变化，使得越来越多的质子进入可见宇宙的范围。

[27]当然，这个假设只能告诉我们为什么N_1、N_2和$\sqrt{N}$的数量级差不多，不能告诉我们为什么它们现在的大小都接近10^{40}。

[28]太阳的发光度正比于G^7，地球公转轨道的半径正比于G^{-1}，所以地球表面的温度正比于$G^{9/4} \propto t^{-9/4}$。

[29] E. Teller,Phys. Rev. 73, 801 (1948).泰勒是匈牙利移民，是一个非常高调的物理学家，在氢弹的研究中起到了关键作用。他和洛斯阿拉莫斯的斯坦·乌拉姆（Stan Ulam）分别想到了引爆核弹的关键方法［苏联的安德烈·萨哈罗夫（Andrei Sakharov）和英国的约翰·沃德（John Ward）也独立地发现了这种方法］。后来，泰勒在罗伯特·奥本海默的庭审中表现得很有争议性，在冷战期间变成了一个极端鹰派份子。在斯坦利·库布里克1964年的黑色喜剧电影《奇爱博士或者我如何学会停止恐惧并爱上炸弹》中，由皮特·塞勒扮演的令人难忘的奇爱博士的原型之一就是泰勒。

[30]改变e的取值不会影响地球的公转轨道，但由于太阳的发光度正比于e^{-6}，所以地表的平均温度正比于 $t^{-3/4}$。如果是这样的话，海洋沸腾时期就离现在非常远，不会影响生物的演化。

[31] J. Cornwell,*Hitler’s Scientists,* Viking, London (2003), pp. 186–190.

[32]这一点首先由罗伯特·迪克在 1957年指出，参见：*Reviews of Modern Physics* 29, 363– 376．这标志着“人择原理”开始应用到宇宙学研究中。更详细的内容参见：J. D. Barrow and F. J. Tipler, *The Anthropic Cosmological Principle*, Oxford University Press, Oxford (1986).

[33] J. D. Barrow,*The Constants of Nature, Jonathan Cape*, London (2003), p. 111.

[34] W. K. Clifford (1876), ‘On the Space Theory of Matter’, in The *World of Mathematics*, Simon and Schuster, New York (1956), p. 568.

[35]罗森是著名的EPR佯谬中的R，那篇论文署名是爱因斯坦、鲍里斯·波多尔斯基（Boris Podolsky）和罗森，1935年发表在*Physical Review* 47, 777。

[36] D. Kennefick, ‘Who’s Afraid of the Referee? Einstein and Gravitational Waves’, http://dafix.uark.edu/~danielk/Physics/Referee.pdf, *and Physics Today* 58 (9), 43-48 (2005).

[37] A. Einstein and N. Rosen, J. Franklin Inst. 223, 43–54 (1937).后

来，人们发现早在 1925年，数学家汉斯·布林克曼就已经发现这种形式的解了：H. W. Brinkmann,*Math. Ann.* 18, 119 (1925).今天，人们称之为pp波（pp wave）。

[38] R. Feynman,*Surely You’re Joking, Mr Feynman!*, Norton, New York (1985).在爱因斯坦方程组中，爱因斯坦张量写作$G_{\fallingdotseq v}$，时空度规张量写作$g_{\fallingdotseq v}$，发音都是“姬—谬—拗”，人们只要一讨论广义相对论，肯定把“姬—谬—拗”挂在嘴边。

第5章
奇异至极的宇宙

有一次，我看到哲学家理查德·洛蒂（Richard Rorty）精神恍惚地站在戴维森食品市场里。他用平静的语调告诉我说，他刚刚在冰冻食品通道看到哥德尔了。

——丽贝卡·戈尔茨坦（美国哲学家、作家）[1]

瑞士奶酪式宇宙

我必须承认，当得知爱因斯坦的电话号码是2807时，我很失望。它真的应当是3622，这样在手机上拼出来就是EMC2（$E=mc^2$）。

——迈克尔·马勒（Michael Mahler）

20世纪 40年代，宇宙学的研究逐渐停滞了下来。持续蔓延的战争迫使物理学家和数学家被重新指派了任务，开展武器、气象学、航空学和密码学的研究。大学不再招收新生，国际上的学术交流也仅限于亲密的盟国之间。爱因斯坦在美国，而其他德国的知名科学家也都逃往英国和美国。宇宙从来没有像现在看起来这么小。

1944年，爱因斯坦在普林斯顿雇了一个新助手。他的助手总是那种有天赋的年轻数学家，这样可以弥补爱因斯坦自知在这个方面的不足。恩斯特·施特劳斯（Ernst Straus，1922～1983）可以说是一个数学神童，他五岁的时候就会用一种有意思的简便方法计算数列，这种技巧能让你在几秒钟内就心算出从1加到100等于多少。[2]他1922年出生在慕尼黑，但在1933年纳粹上台后，全家逃难到巴勒斯坦。他在那儿读了高中，然后进入耶路撒冷的希伯来大学。施特劳斯没有拿到本科学位，1941年，不到二十岁的他就跑到纽约的哥伦比亚大学开始读起了研究生。1944年，他被普林斯顿高等研究院的爱因斯坦招为了新的研究助理。[3]



图5.1 恩斯特·施特劳斯

年轻的施特劳斯没有多少物理学基础，他的数学能力也主要体现在数论和理论数学方面。但随着纳森·罗森（1935～1945年间合作）和利奥波德·因费尔德（Leopold Infeld，1936～1938年间合作）的离开，他必须及时补上这个空缺。1945年春天，爱因斯坦和施特劳斯求解爱因斯坦方程组后，发现了一个新的宇宙模型。[4]这种模型同弗里德曼和勒梅特的简单的膨胀宇宙模型非常像，只包含无压强的物质（例如星系）。但是这种宇宙的空间中含有球形的空腔，就像一块瑞士奶酪里的空泡（图5.2）。每一个空腔的中心都有一定质量，质量的大小等于形成这

个空腔所要挖走的物质质量。这种模型向更现实的宇宙跨出了一步，其中的物质并不是均匀地分布在四面八方，而是聚集成团，就像星系那样，在空间中散播开来。

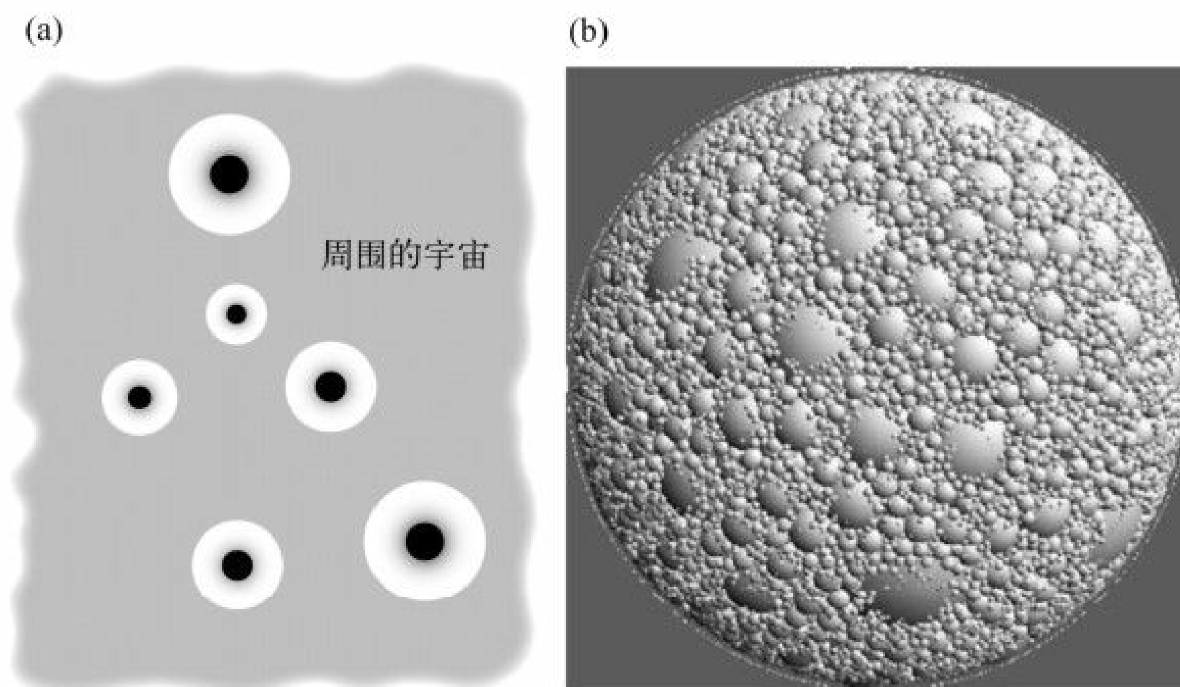


图5.2 (a)瑞士奶酪式宇宙模型。球状区域所有的质量都集中在了中心处，其他地方是空的。(b)多伦多大学阿伦·阿塔德（Allen Attard）用计算机模拟的瑞士奶酪式宇宙。开始时宇宙中的物质是均匀分布的，然后34 667个大小不等的球状空间就产生了，其中的物质都被压缩到了中心的一点上；这些孔洞占据了80%的空间

每一个“空腔”都是球状，在特定的初始条件下，这种新的瑞士奶酪式宇宙就能和托尔曼的不均匀宇宙联系起来。同往常一样，想要从爱因斯坦那复杂艰涩的方程组中找到一个精确解，就必须先对解的形式进行某种简化，这样代入方程以后才算得动。然而，这就像喜剧演员格劳乔·马克思（Groucho Marx）在回忆录中所说的那句著名悖论，他不愿加入任何同意他加入的俱乐部。[5]类似地，凡是能从爱因斯坦方程组得到的简单解，总会有一些特殊性质，让它变得很反常、没意思。

爱因斯坦和施特劳斯的这个解很简单。由于是球对称的，所以就不会产生任何引力波——不像爱因斯坦和罗森的那个柱状宇宙模型。这让人不禁想问，如果用某种方法把所有的反常都整合在一起，会得到什么结果呢？如果所有反常性质一起出现的话，求解爱因斯坦方程组当然就

是痴人说梦。不过，有一种方法确实能让人隐约看到这种宇宙是什么样子的。

被扰动的宇宙

“有些树，华生，它们长到一定高度后，就突然开始长歪了。”

——阿瑟·柯南·道尔（1859～1930）[6]

自从20世纪20年代弗里德曼做出膨胀宇宙模型的惊人发现以后，苏联的宇宙学就没得到多少发展。苏联人认为，爱因斯坦的理论肯定跟与辩证唯物主义相对立的唯心主义有着千丝万缕的联系。[7]而且当时的政治气氛让苏联人人自危。该国最伟大的理论物理学家列夫·朗道（Lev Landau，1908～1968）在1938年就曾被判入狱。1939年，在其他国际知名的物理学家如彼得·卡皮查（Peter Kapitsa，1894～1984）和尼尔斯·玻尔等人的多方努力下，朗道才得以重获自由。

自1932年起，年仅24岁的朗道就当上了新成立的乌克兰哈尔科夫物理技术研究所的主任。从青少年时期开始，朗道在物理学方面就有了极高的造诣。1962年，由于他在低温液氦的超流理论中所作的贡献，朗道获得了诺贝尔物理学奖。[8]在新成立的研究所中，朗道的鼎鼎大名吸引来了许多优秀学子，并且他还为高等物理学教育制定了新的标准。他最著名的学生是哈尔科夫的一名年轻人，伊夫金·栗弗席兹（Evgeny Lifshitz，1915～1985）。栗弗席兹18岁时就拿到了物理学和力学的学位，又花了一年多一点儿的时间挑战了朗道大师的课程，全部通过，并获得了博士学位。[9]几年以后，栗弗席兹就可以和朗道在一系列物理学的高级教材中共同署名了，“朗道和栗弗席兹”的名号对全世界的物理学家来说都是如雷贯耳，他们以优雅、简约、统一的风格向读者解读了理论物理学的核心问题。

朗道的被捕对栗弗席兹来说是一个潜在的威胁，因为他和他的导师关系密切。[10]幸运的是，朗道出狱的时候政治气氛有所缓和，于是他和栗弗席兹搬到了莫斯科的物理问题研究所。然而，很快战争就爆发了，一切日常工作也随之中断。1941年9月，莫斯科之战打响，一直持续到了1942年1月。由于莫斯科濒临崩溃，加上饥荒的蔓延，许多人被迫撤离，或者向东逃难。许多物理学家被安置到了喀山。但恶劣的天气挽救了莫斯科，倾盆大雨和刺骨的寒冷让德国侵略军措手不及，停止了推进。就是在这样艰苦卓绝的条件下，栗弗席兹开始了新的宇宙学研究。就像在他之前的弗里德曼，栗弗席兹的着眼点也是在数学方面，而不是天文学。

我们已经看到，科学家们开始对各向异性、非均匀的宇宙产生兴趣，例如弗里德曼、勒梅特和德希特发现的宇宙模型。先是卡斯纳和托尔曼，然后是爱因斯坦、罗森和施特劳斯，都开始着手研究各向异性、非均匀的宇宙。研究这种宇宙模型明显要更有现实意义，因为真实的宇

宙并不是完全光滑、各向同性的。

栗弗席兹用一种物理学中很普通的方法来研究这个问题：先求得一个简单的精确解，然后对它进行小幅度的扰动，看看结果会如何变化。方程组会如何限定这些扰动的演化行为呢？随着时间的流逝，它们是会逐渐消失，还是会越变越大？如果这些扰动逐渐消失，说明这种简单的宇宙模型在微扰下是稳定的，不会演化成另外一种结构。然而，如果小的扰动会随着时间增大，这种简单的宇宙模型就是不稳定的，将来会逐渐演化到另外一种状态。

1946年，栗弗席兹的一篇论文发表在苏联的学术期刊上，文中研究了弗里德曼的各向同性、均匀宇宙在微扰下的演化行为。[11]他似乎只是把它当作纯数学问题来研究，而没有同“宇宙中为什么会存在星系这样的不均匀结构”的问题联系起来。他证明，这种宇宙中可以存在三种不规则性：第一种就是简单地让物质密度的大小随地点变化，第二种是让物质缓慢地转动，而第三种就是向光滑的空间中引入微小的引力波涟漪。如果这三种不规则性的幅度很小，它们之间就不会相互影响，乱作一团；但如果幅度很大，情况就相反了。小幅度扰动的影响相互独立，很容易单独分析。以一个圆球为例，第一种微扰会改变圆球的密度和形状，第二种会让圆球自转，第三种则会把圆球压扁，变成椭球形，但体积没有变化。

栗弗席兹证明，密度分布的微小差异会随时间的推移而变得越来越明显。其实早在17世纪时牛顿就已经知道了，他在给剑桥三一学院长期掌权的（专制的）理查德·本特利（Richard Bentley）的信中说，一旦向完全均匀分布的物质中引入微扰，高密度区域吸引的物质就会越来越多，而稀薄的区域就会变得越来越稀薄，物质的分布于是就越来越不规则。我们把这种过程叫作“引力不稳定性”（图5.3）。在不膨胀的空间里，物质分布的这种不稳定性会很快体现出来。栗弗席兹证明，在膨胀的宇宙之中，这种不稳定性也会体现出来，但不规则性的积累过程会比较慢，因为物质粒子在结团的同时需要克服膨胀所产生的拉力。他同时也证明了，如果物质聚集时存在阻力，结团的尺寸又足够小的话，不规则性的积累最终会停下来。这个问题从总体上看很简单。如果宇宙开始时接近各向同性、均匀的状态，那么随着宇宙的膨胀，密度的不均匀就会越来越明显。[12]这是第一次有人通过计算得出，各向同性、均匀的宇宙其实很特殊。经过一百多亿年的膨胀，宇宙中的不均匀性还是如此不起眼，说明140亿年前，宇宙初始的不均匀性更是微乎其微。

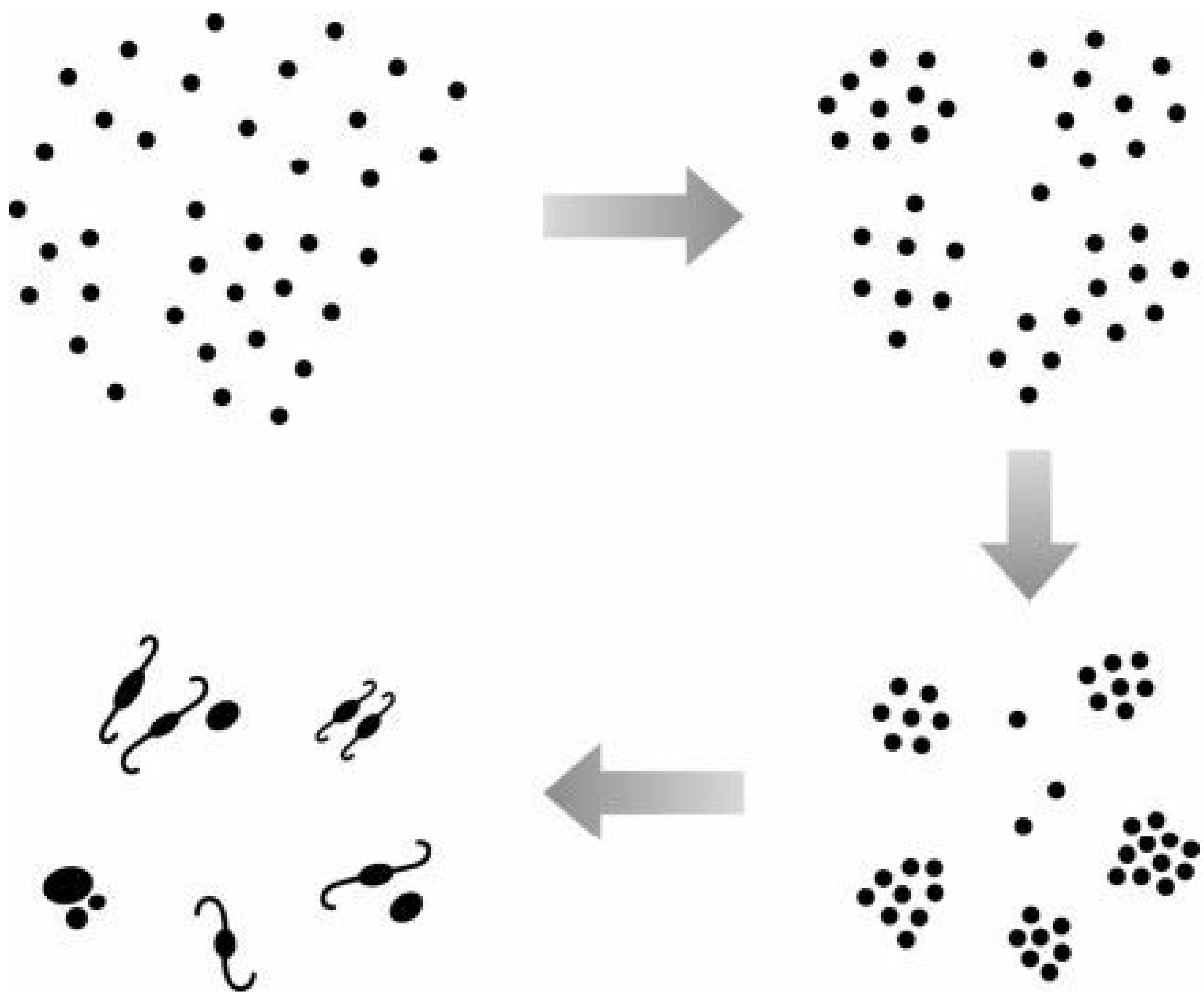


图5.3 在引力的作用下，如果物质的分布稍微有点儿不均匀，就会形成团块的结构。密度大的区域会产生更强大的引力，吸引更多的物质，使其他区域变得更加稀薄

渐渐地，其他天文学家也开始进行相关计算，试图解释宇宙为什么会形成像星系这样的结构。不幸的是，这些解释都不怎么有说服力。为了让不规则性刚好在晚近的宇宙历史中形成星系结构，你必须在宇宙刚开始膨胀时就给不规则性选定合适的大小。但没有合理的理由能让人随意设置宇宙初始不规则性的大小，于是这并不是一个全面的理论。

至于栗弗席兹算过的其他不规则性，你会发现自转物质形成的巨大漩涡会在宇宙膨胀时转得更慢。后来，一些宇宙学家认为，这说明很久以前的漩涡转得更快，也就是说，我们今天看到的所有旋转的星系都是由一个原始的混沌漩涡形成的。最后，爱因斯坦和罗森纠结不清的引力波，也能以微扰的形式存在于一个绝对光滑的宇宙中。栗弗席兹的计算表明，这些微扰都会随着宇宙的膨胀而衰减，同时以波的形式四处传

播。

栗弗席兹虽然只是把这些计算当作一种数学练习，但这却是宇宙学研究的一个里程碑。从那时起，这些计算就以各种方法重复了一遍又一遍。如今，我们在观测宇宙的微波背景辐射和星系群的结构时，就会看到这三种微扰留下的各种影响，整本整本的众多专著都在讨论这些影响。总之，栗弗席兹的计算给我们提供了一种处理微扰的方法，以探究原始宇宙的微小扰动是如何随时间而发展变化的。

薛定谔的宇宙

尚未观察到的任何东西都是波，已经观察到的任何东西都是粒子。

——威廉·劳伦斯·布拉格（William Lawrence Bragg，1890～1971，1915年诺贝尔物理学奖得主）[13]

当爱因斯坦的理论颠覆我们对宇宙的认识时，在原子尺度上，一场更迅猛的变革正在改变我们对物质的认识。量子力学是一种认识物质和光的行为的新理论，它的出现要大大归功于爱因斯坦，但更要归功于尼尔斯·玻尔、沃纳·海森堡、保罗·狄拉克、马克斯·玻恩、沃尔夫冈·泡利（1900～1958）和埃尔温·薛定谔（1887～1961）等人。在量子力学的指引下，人们逐步揭示了分子和原子的结构，解释了化学元素周期表的规律，并预言了固体材料的大量性质。但是量子力学和宇宙学没什么交集。狄拉克虽然在宇宙学领域跨了次界，但他提出引力的强度有可能随宇宙年龄的增长而变慢时，并没有用到量子力学。勒梅特曾经有个想法，想知道宇宙如何从某种超级致密的原初状态诞生，他将之取名为“原始原子”。为了研究其中的奥秘，必须把量子力学和爱因斯坦的引力理论结合起来，但勒梅特不知道该怎么做。有的物理学家想要搞清楚，描述物质的诞生是不是要靠量子力学，而理查德·托尔曼给出一个模糊的推测：星系中心就是物质和辐射源源不断地进行相互转化的地方。这些思考都没什么结果，但很快，一位著名的物理学家第一次把目光投向了量子的宇宙。

埃尔温·薛定谔在1926年发现的薛定谔方程，是数学物理领域中最重要的方程。薛定谔方程的解能够描述所有分子和原子的结构，能够解释所有材料科学和化学现象。薛定谔的父母很有钱，只有他这一个儿子，看到儿子极具天赋，就给他请了私人教师。他是维也纳大学的优秀毕业生，随即在那里任教，后来又在苏黎世、柏林、剑桥和都柏林得到了教授职位。他和保罗·狄拉克分享了1933年的诺贝尔物理学奖。

薛定谔的个人生活很另类。在他1933年在牛津大学曼达琳学院短暂停留期间以及次年普林斯顿大学邀请他担任教授的事情上，他希望与妻子和情妇共居一处的态度着实引发了争议；此外，他们三人还陆续跟其他物理学家和物理学家的太太们搞过暧昧。最后，薛定谔在1936年搬到了奥地利的格拉茨。1940 年，他受到爱尔兰总理埃蒙·德·瓦莱拉

（Éamon de Valera）的邀请，要他赴都柏林加入都柏林高等研究院，这所研究院模仿了为爱因斯坦而建立的普林斯顿高等研究院。他接受了邀请，但后来他在都柏林与有夫之妇搞婚外情，还弄出了几个私生子，这些丑闻给他的职业生涯留下了不少污点。如果说薛定谔30年代在牛津的生活过于放荡，那么50年代在都柏林的生活简直是无比放荡。1961年，薛定谔去世，埋在了奥地利蒂罗尔州的阿尔卑巴赫村，远处就是他深爱的阿尔卑斯山。他的墓碑上则刻着薛定谔方程（图5.4）。



图 5.4 薛定谔的墓地位于奥地利的阿尔卑巴赫村，墓碑上面刻的是薛定谔方程

薛定谔对爱因斯坦的广义相对论预言的宇宙膨胀产生了浓厚的兴趣。他在宇宙学研究中的首次试水，就是想知道量子力学能在宇宙学中预言什么。他决定先研究膨胀宇宙中的波动行为。[14]结果不但深入了解了通常的声波和光波在膨胀宇宙中的行为，更阐述了量子的波动的传播机制，量子的波动能告诉我们观察到特定的物理事件的概率。[15]

薛定谔做出的重要发现在当时无人喝彩（包括他自己也未意识到其重要性）。宇宙膨胀的过程会把其中的量子真空①的能量转化为真实存在的、可测量的粒子。如果宇宙不膨胀，真空中会有成对的粒子和反粒子不断产生，然后又湮灭成辐射。在这个过程中能量是守恒的，而这个沸腾的景象描述的就是量子真空。然而，如果真空所在的空间膨胀得非常快，或是处在一个极度不均匀的引力场的作用下，那么真空中的粒子和反粒子就会被不同的作用力拉开，无法再湮灭成辐射。因此，真实的、探测得到的粒子和反粒子出现了，驱动了宇宙膨胀的能量是它们的山珍海味，引力场不均匀的分布是它们的玉露琼浆。20世纪70年代，史蒂芬·霍金做出了一个举世震惊的发现：就像膨胀的宇宙，黑洞的边界也会不断产生正反粒子对，最终导致了黑洞总质量的减少。[16]从那以后，物理学家们又对膨胀宇宙产生粒子对的惊人现象做了更深层次的研究。

① 在量子力学中，真空指的是量子场能量（局部的或全局的）最低的状态。真空中的量子场会不断产生和湮灭虚的粒子对，这些虚粒子对会产生很多可观测的效应，例如改变原子的光谱。因此，真空不是真的“空”。人们推测真空中蕴涵的能量会影响宇宙的演化。——译者注

当时，薛定谔并不觉得这个过程对膨胀的宇宙有多重要，因为在今天的宇宙中，这种效应太微弱了，不影响宇宙的演化行为。然而，在宇宙膨胀的最初阶段，膨胀的速率高得惊人，辐射的能量密度也要比今天高出近10^{128}倍，粒子对产生的过程就不能忽略了。宇宙现在的某些特点可能就是这个过程所导致的。20世纪70年代，有人提出宇宙膨胀中的许多不对称和不均匀的特性都被粒子对产生的过程完全减弱了，抹平了，因为这种过程能弱化不同位置、不同方向间的能量差异，让它们渐渐地消失。

然而，1939年的宇宙学家们还没有做好迎接量子力学的准备。要是当时物理学家们乐于接受量子力学的话，恐怕结果就会大不相同。但是

爱因斯坦当时坚决反对量子理论（“上帝不扔骰子”是他的名言），并且强烈抵触将量子理论应用于宇宙整体的做法。不过，粒子对产生的理论不久就变得极富争议性，对此我们将在下一章看到。

哥德尔的旋转宇宙

神学世界观认为，世界和万物都有一个积极明确的意义。既然我们的世俗存在从本质上讲意义就不确定，由此可以直接得出，这只不过是走向另一种存在的手段。世界万物都有一个意义的想法类似于世界万物都有一个原因的理念，后者是所有科学理论的根基。

——库尔特·哥德尔（1906～1978）[17]

爱因斯坦上了年纪以后力不从心，不再解决老问题，也不提出新问题了。他喜欢告诉人们说，他去办公室上班，“只是为了获得能和库尔特·哥德尔一起步行回家的荣幸”。哥德尔是 20世纪最伟大的数学家之一，是自亚里士多德以来最重要的逻辑学家。1906年，他出生在奥匈帝国的布尔诺，也就是爱因斯坦发表他的狭义相对论、布朗运动和光电效应论文的下一年。1924年，哥德尔进入维也纳大学学习物理学，但很快又被数学吸引住了。哥德尔给他的教授留下了深刻印象，于是很快被邀请加入了鼎鼎大名的研讨小组“维也纳学派”（the Vienna Circle），他们在咖啡馆见面，讨论哲学、逻辑学和科学问题，其中经常出席的有路德维希·维特根斯坦、伯特兰·罗素和卡尔·波普尔。但哥德尔是其中的另类：他是唯一不相信经验是知识的唯一来源的人，也不相信数理逻辑是解决哲学问题的唯一工具。

哥德尔最出名的成就是证明算术的不完备定理。这个定理是说，任何一个允许定义自然数的逻辑体系，总是包含这样一些命题，既不能用系统内部的公理证明为真，也不能证明为假。这引起了数学家和哲学家的一片哗然。从这个定理会得到许多令人意想不到的推论，例如，没有一种电脑程序，能不通过篡改操作系统，就可以检测那些篡改操作系统的程序。因此，没有一种不干扰、不篡改操作系统的杀毒软件能找到你电脑上的所有病毒。

和爱因斯坦一样，为了逃避自己祖国日益猖獗的法西斯势力，哥德尔后来去了普林斯顿高等研究院。1933～1934年间，哥德尔第一次访问了这个研究院，但后来自从他患上某种精神障碍以后[18]，就再也没有来过。在大学外遭到了某些纳粹分子的袭击之后（很可能把他当作了犹太人），1939 年秋天，哥德尔终于和妻子一起离开了维也纳。可是他选择了去普林斯顿最远的一条路，沿着横贯西伯利亚的铁路穿过亚洲，乘船从日本到旧金山，到时已经是1940年3月了，最后乘火车横穿美国到达普林斯顿。

1942年，哥德尔成了爱因斯坦的亲密伙伴。两个人都有相似的文化背景和对哲学的强烈兴趣，这跟他们周围的美国物理学家大不相同，而且他们都可以讲德语。

有好几次，哥德尔答应爱因斯坦把他研究相对论的心得记下来，但直到1947年，激动人心的事情才终于发生。当时哥德尔在给他母亲的信中谈到，他已经加入到了相对论的研究工作中，到了1947年夏天的时候已经发现了一些不同寻常的东西。出乎所有人的意料，他一直在寻找新的爱因斯坦方程组的解。结果就连爱因斯坦都大吃一惊。

哥德尔的宇宙是一个不断旋转的宇宙（图 5.5）。这种宇宙不膨胀，所有的物质都绕着一个对称轴匀速转动。其中也包含了爱因斯坦的宇宙学常数，但不同的是，这里的宇宙学常数小于零，因此产生的是引力，和物质的引力一起抵消了转动产生的离心力。这本身就够有趣的了，但哥德尔的宇宙还有一个完全令人无法想象的性质：它允许时间旅行。哥德尔证明，时空中的一些路径形成了闭合的回路。大多数人，包括爱因斯坦，都相信这种事情应该违背了其他的物理定律，并且会导致科幻电影里经常演到的逻辑悖论（例如，杀死婴儿时期的自己）。[19]但其实爱因斯坦的理论是允许时间旅行的，而且不与任何已知的自然法则相矛盾。物理学家弗里曼·戴森（Freeman Dyson）回忆起他第一次见到哥德尔时，所听到的关于时间旅行的点点滴滴。那是在1948年，当时他还是个刚到普林斯顿的年轻人。

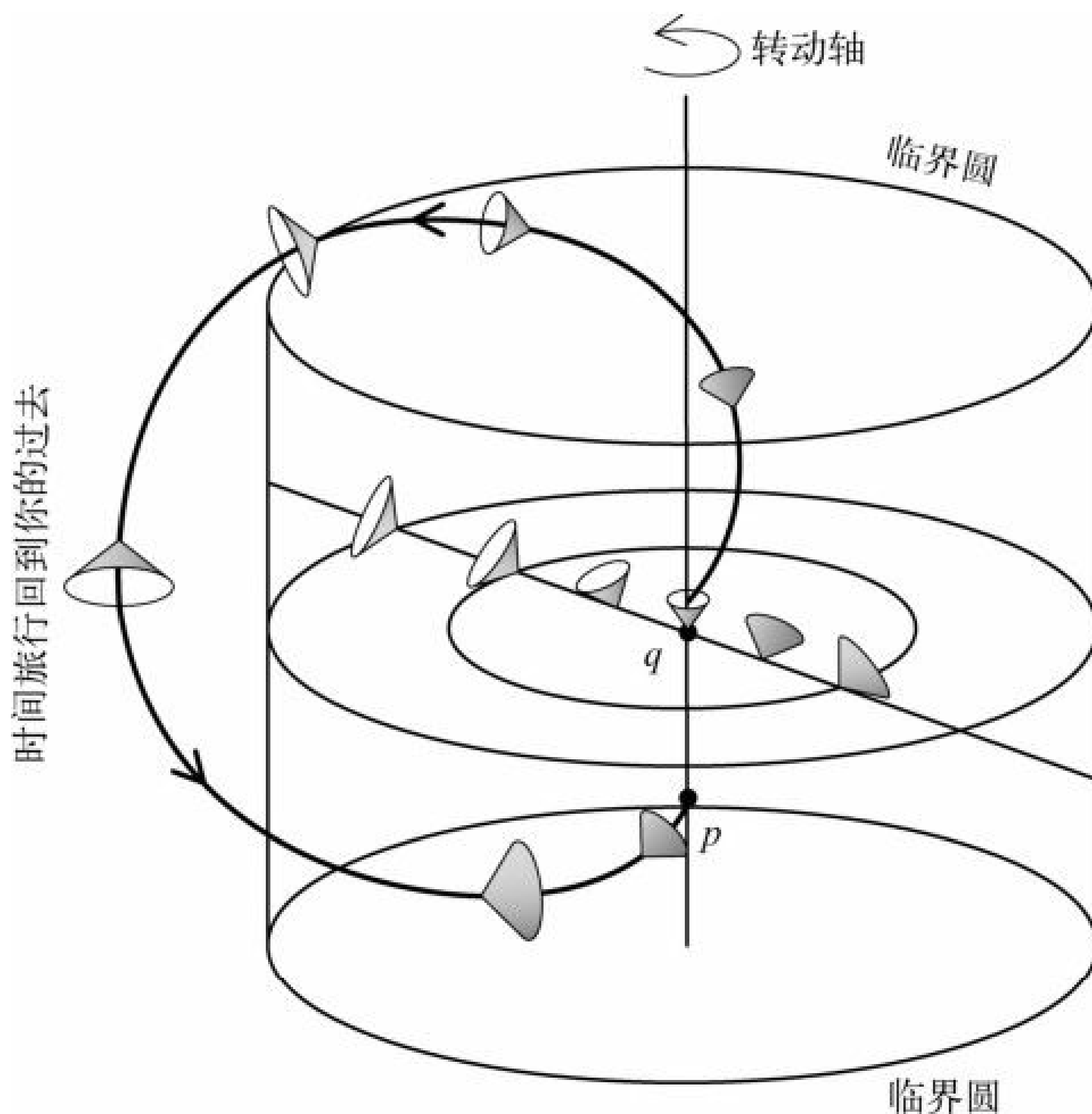


图5.5 哥德尔的宇宙。物质沿着中心轴匀速转动，转动对光锥造成了影响。光锥表示每一点所发出的光线所经过的地方。[①]当我们离开中心时，光锥就开始倾斜，并扩大了开口，这是因为转动的线速度增大了。在距转动轴一定距离的地方，光锥完全翻倒，与空间相切，然后倒扣了起来，于是光线就沿着开口朝下（过去）运动。假设你的星球以前在p点，现在在q点。要想再次回到p点的话，你就要朝着临界圆外部的一个点加速运动，然后向过去运动到p点之前的某个地方，进入那时的临界圆，然后再向未来运动到p点。你总是在走向“你的”未来，但却回到了你的过去

① 任何物体的运动速度都不会超过光速。因此，图中每一时空点上的（未来）光锥代表彼时彼处的物体所能影响到的未来事物所处的时空区域。图中外侧的某个（未来）光锥的开口与中心处的光锥开口相反，可理解为前者的“未来”指向了后者的过去。——译者注

那是在1948年9月。我是普林斯顿高等研究院的一名年轻的新成员。让我惊讶的是，我第一个见到的人竟然就是库尔特·哥德尔本人，更让我受宠若惊的是，他邀请我去他家。但不管怎样，我觉得自己好幸运啊。我发现他特别亲切，善于交际，不像我想的那么孤傲。而且身体还很健康！所以，他请我去他家，我们边走边聊一些物理问题。看起来他对物理学的了解很深入，而且这些研究都是他自己搞的。前几年他听从爱因斯坦的建议，研究了旋转宇宙模型，这就是他的科研工作。让我有点儿吃惊的是，一方面，他绝对是一个出类拔萃的数学家，着实撼动了数学大厦的基础……而这样的人要做的是……让人不解的是，他会做一些相对来讲微不足道的事情，比如证明旋转宇宙的存在。当然，在物理学中这都不是什么有趣的问题。对，他自己相当清楚，他并不是不懂物理学，他知道这确实不是物理学的主流。但无论如何，他就是这样。当然，后来我们继续见面。见面时，他经常会问我：“他们发现了吗？他们知道宇宙在没在旋转了吗？”他认为这种事情是可以被观测检验的，于是我不得不说明，现在的观测水平离这个问题的要求还差十万八千里，他总是很失望。他给我打电话的时候，常常会问：“他们发现了吗？”而我总是不得不告诉他还没呢。[20]

现在看来，哥德尔的发现[21]对我们的宇宙学研究特别重要。这个模型告诉我们，宇宙可能有一些特殊的总体性质，从局部是看不出来的。只是因为时间和空间在太阳系中看来很正常，并不代表它们不会



图5.6 爱因斯坦和哥德尔在普林斯顿

在整个宇宙的尺度上以奇怪的方式纠缠在一起。尽管哥德尔的宇宙并不像我们的宇宙那样在膨胀，但其表现出来的时间旅行的特性也有可能在其他的、像我们宇宙一样膨胀的旋转宇宙中出现。

起初，一些著名物理学家质疑哥德尔的宇宙允许时间旅行，但实际上他们是把这些时间旅行的历史的特性理解错了，最终哥德尔的推导被证明是正确的。在哥德尔的一些笔记中，他好像觉得用时间循环来证明永生的可能性是件相当有趣的事情——有一次，一位同事发现他在黑板上倒着写字，表演时间旅行。不巧的是，要想在哥德尔的宇宙中进行时间旅行，就必须接近光速，并且要求物质以一种不同寻常的方式分布其中。对太空漫游者来说，这不是个靠谱的提议。我们还得注意的是，哥德尔的宇宙并不能证明电影《回到未来》的情节可以实现。你不可能改变历史。就像诗人萨穆埃尔·巴特勒（Samuel Butler，1613～1680）所说的，即使是上帝也无法改变历史——只有历史学家才做得到。

假设我背会了莎士比亚的悲剧《麦克白》，穿越历史回到过去，见到了年轻的莎士比亚，他那时还没开始写剧本。我把《麦克白》的原文和情节给他详细地讲述一遍。莎士比亚记下了每一个字，把它写了出来，然后出版了《麦克白》。那么《麦克白》到底是怎么来的？我从莎士比亚那儿学来，而他又从我这儿学来。这其中没有开端，它只是存在着。

关于“如果我杀死了祖母”的逻辑悖论有各种各样的版本，关心时间旅行问题的哲学家把它们归纳为“祖母悖论”。[22]这种悖论推翻了各种时光旅行的假说。[23]自从1895年赫伯特·乔治·威尔斯（1866～1946）的经典科幻小说《时间机器》首次提出这个想法以来，时间旅行就成为了科幻小说的重要题材。

我们该不该承认，这些“改变历史”的例子说明时间旅行的想法存在某种根本缺陷？不。改变历史的想法中确实有一些自相矛盾的地方。历史就是过去的事。你不可能既改变了历史，又保留你现在的经历。不可能有两个历史。如果你能回到过去，设法阻止了自己的出生，那么就不可能有机会回到过去，阻止自己的出生。

通常我们认为时间的流逝像一条直线。在时间旅行之中，这条线闭合成一个圆（图 5.7）。想象一条直线上，有人排队走过。谁在前，谁在后是很显然的事情。这就像线性的时间：你总可以毫不费力地分清一件事发生在过去，还是在将来。

现在，我们假设这条线上的人们绕成了一个圆。从局部看，有的人在前，有的人在后，但是总体上看，在整个圆上根本分不清“前面”和“后面”——每个人都既在其他人的前面，又在其他人的后面。所以就不能说谁在谁的前面，谁在谁的后面了，而应该说前后都在。[24]

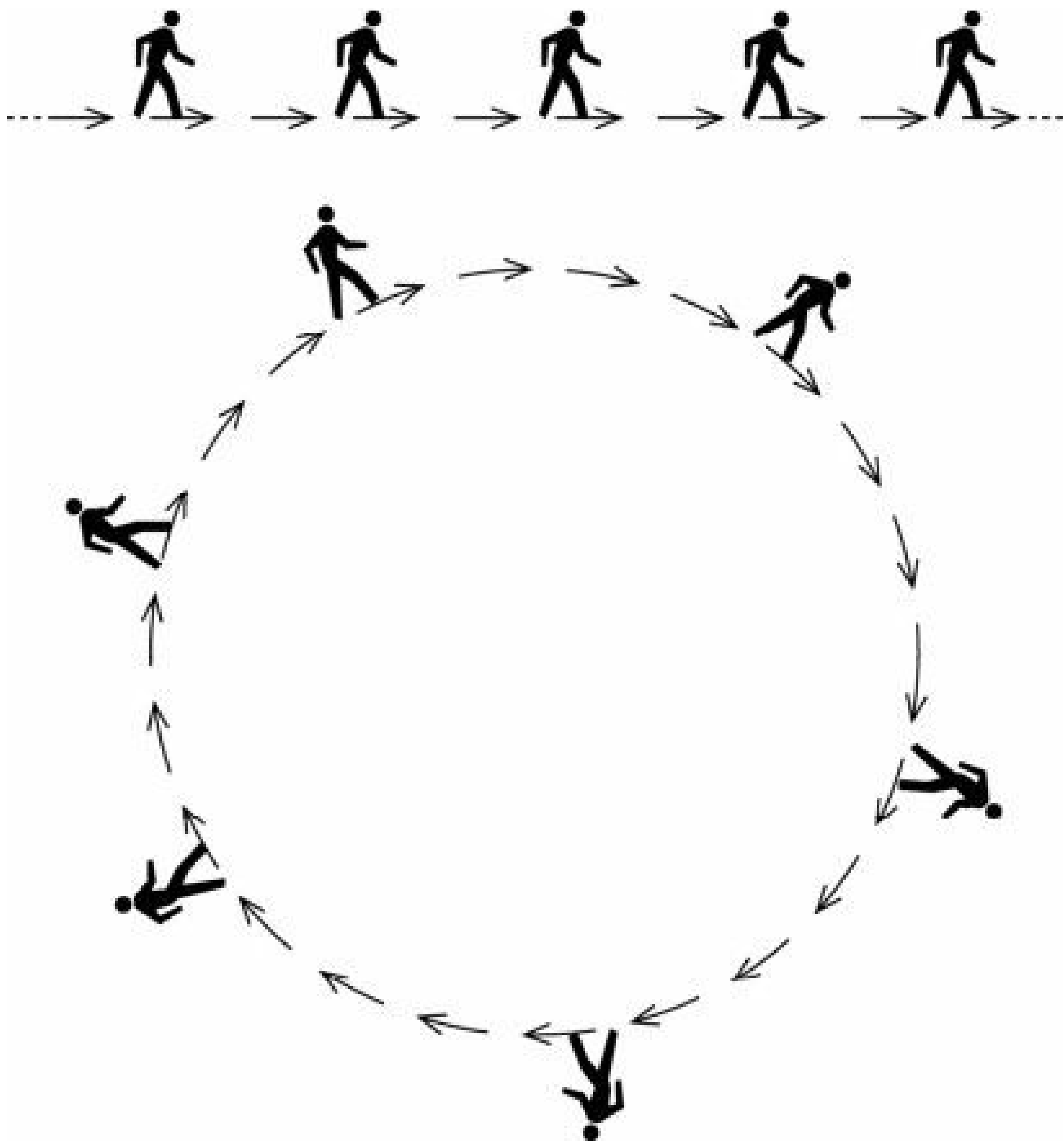


图5.7 路人沿着直线和圆一个接一个地排列。对于直线上的人来说，他要么在某个人的前面，要么在某个人的后面。对于圆上的人来说，他既在某个人的前面，又在某个人的后面

并且时间旅行也是这种情况，其中并没有绝对的过去与未来。循环的时间里分布着一系列逻辑自洽的事件。当下的就是当下，曾经的就是曾经。你可以成为过去的一部分，但你无法改变过去。如果你活得足够长，你会一遍又一遍地经历相同的事情，循环往复。

这里有一个自洽的时间旅行的历史的例子。想象一下，你回到过去，准备好要朝婴儿时的自己开枪。你决心要在宇宙中创造出一个悖论。你瞄准了你母亲怀抱中的自己。当你正要扣动扳机时，由于婴儿时期的自己从母亲的怀中摔了下来，你肩膀上形成的旧伤突然让你的胳膊发出一阵痉挛，结果导致你射偏了。然而，枪声足以吓到你的母亲，把小宝宝摔落在地上，摔伤了肩膀。历史自洽了，宇宙安全了，历史学家们放心了。

指引哥德尔发现旋转宇宙模型的是他的一个信念，他想证明时间的流逝并不是客观的：关于时间并没有绝对的标准。他的宇宙很古怪，因为从其中的每一点看，宇宙都是一样的（只有匀速转动，没有膨胀），但是没有额外的标准可用以衡量旋转的速率（因为他理论中的宇宙只有一个，不存在“宇宙外部”）。哥德尔小心地计算出，进行循环式时间旅行需要的速度大小，并花了许多时间用来收集天空中星系分布的数据，因为他相信宇宙在旋转。唉，可惜我们的宇宙不是哥德尔的宇宙。宇宙正在膨胀，如果它开始旋转的话，它的膨胀速度就会变得特别慢。这很容易验证，因为微波背景辐射从四面八方传来，其密度会随不同的方向而变化，幅度不超过十万分之一。[25]就像地球自转会导致形状变扁，宇宙的旋转也会扭曲微波背景辐射的温度谱形状，使得它沿着自转轴的方向最热，垂直于自转轴的方向最冷。

虽然哥德尔的发现实际上并不能描述我们的膨胀宇宙，但它给我们带来了一些期待，爱因斯坦的方程组中可能还隐藏了更多类似的奇妙事物。尽管局部性质完全正常，宇宙还是会有一些奇特的总体性质。哥德尔的宇宙表明，宇宙的旋转以一种极端的方式扭曲了空间，以至于把时间都闭合了。哥德尔证明，这样的宇宙满足爱因斯坦场方程，但不满足牛顿引力。

遗憾的是，后来哥德尔就再也没有发表宇宙学的研究工作了。他的注意力转向了逻辑学和哲学中最最艰涩的问题。宇宙学家们花了大量时间，想要了解他是如何找到这个解的，但哥德尔没有留下任何线索。真是个科学怪人。

注释

[1] R. Goldstein, *Incompleteness: The Proof and Paradox of Kurt Gödel*, W. W. Norton, New York (2005).

[2] 1加100，2加99都等于101。从1加到100，总共得到50个101，因此结果是50×101=5050。同理，从1加到N的结果是N(NǴ1)/2。有个著名的故事是，高斯九岁的时候，老师想让学生花很多时间来计算从1加到100，结果高斯一下子就想到这个方法了。

[3] 有意思的是，施特劳斯把期间自己与爱因斯坦的往来信件保留

了下来，包括32封书信和15份手稿。他的家人于2006年6月在伦敦奥林匹亚公园的古籍展上拍卖了这些文稿，底价是150万美元，参见：http://www.guardian.co.uk/uk/2006/may/22/science.research.

[4] A. Einstein and E. G. Straus, ‘The Influence of the Expansion of Space on the Gravitation Fields Surrounding the Individual Stars’, Re. Mod. Physics 17, 120 (1945), and 18, 148 (1946).第二篇文章是对第一篇文章的补充和修正。在1946年和1949年，施特劳斯又和爱因斯坦合作了两篇论文，爱因斯坦试图创建一个比广义相对论更深刻的统一场论。后来，施特劳斯所有的论文都是关于数论的。

[5] 这是马克思给比佛利山庄的弗雷厄俱乐部发的电报内容，引自：*Groucho and Me*, Da Capo Press, New York (1959), p. 321.

[6] ‘ The Adventure of the Empty House’, in A. Conan Doyle, *The Return of Sherlock Holmes*(1903).福尔摩斯在前面的故事中被认为已经在赖兴巴赫瀑布牺牲了，害死他的是莫里亚蒂教授。这是其后福尔摩斯探案的首部续集。

[7] 弗拉基米尔·福克（1898～1974）是当时首屈一指的科学家。在弗拉基米尔·福克的学术生涯中，曾作为学生聆听过弗里德曼的讲座课程。随后，他进入相对论研究领域，试图在意识形态的重重阻挠中觅得出路。关于这段经历的详细描述，参见：G. Gorelik, ‘Vladimir Fock: Philosophy of Gravity and Gravity of Philosophy’, in *The Attraction of Gravitation*, Birkhäuser, Boston (1993), pp. 308–331.

[8] 遗憾的是，那一年他遭遇了一场严重的车祸，整整昏迷了六个星期。后来他一直活到了1968年，但再也没能重新找回从前的天赋。朗道对物理学每个方向的细枝末节都了如指掌，而且他在任何方向上都能够做出新的发现。

[9] 一共只有43个学生通过了资格考试，得以同朗道一起做研究。栗弗席兹是其中通过速度最快的。

[10]他必须避免在错误的场合引起人们的注意。为此在1938～1939年间，他在莫斯科和哈尔科夫换了好几个不靠谱的工作，然后有三个月的时间完全没有工作，只是待在遥远的克里米亚。

[11] E. M. Lifshitz, J. Phys. (USSR) 10, 116 (1946).

[12]密度不规则性的演化行为，就好比诸多闭合的、高密度的弗里德曼宇宙被嵌入到了一个平坦的、体积无穷大的宇宙中。这些宇宙会膨胀得更慢，因为它们当中包含了更多物质。因此，这些闭合宇宙和平坦背景之间的密度差异会逐渐扩大。

[13]弗里曼·戴森转述，参见：J. D. Barrow, P. C. W. Davies and C. L. Harper, (eds.),Science *and Ultimate Reality*, Cambridge University Press,

Cambridge (2004), p. 83.

[14] E. Schrödinger, ‘The Proper Vibrations of the Expanding Universe’, *Physica* 6, 899–912 (1939).后来，他将他的成果推广到了描述电子行为的狄拉克方程，参见：*Proc. Roy. Irish Acad. A* 46, 25–47 (1940).他的两本书里对此有详细介绍，参见：*Spacetime Structure*, Cambridge University Press, Cambridge (1950), and *Expanding Universes*, Cambridge University Press, Cambridge (1957).

[15]薛定谔从未全盘接受薛定谔方程中波函数的标准解释。他坚持自己的看法，认为波函

数表示某种类型的荷密度，而不是玻恩所提出的表示某种测量结果出现的几率。 [16] S. W. Hawking, ‘Black Hole Explosions?’,*Nature* 248, 30 (1974).

[17]见于1961年10月哥德尔写给他母亲的信。

[18]他对许多东西都显得神经兮兮的，总觉得有人要毒害他，所以几乎什么都不吃。1978年，哥德尔去世时只有36公斤，而且死因跟饥饿有关。他的妻子阿黛勒既要当厨师，又要替他试吃，还要当保姆，于是1970年阿黛勒去世后，哥德尔的情况明显恶化。

[19]早些时候，人们发现爱因斯坦方程组存在一个包含无压强物质的柱对称旋转解。由于高速的旋转导致时空强烈扭曲，在圆柱的某个范围之外，也允许时间旅行的发生。这个解是由卓越的匈牙利数学物理学家柯涅流斯·蓝佐斯（Cornelius Lanczos，1893～1974，曾在1928～1929年间做过爱因斯坦的助手）在1924年发现的，论文发表在Z. *f. Physik* 21, 73。后来这个解又在1937年被荷兰数学家威廉·范斯托克姆（Willem van Stockum，1910～1944）重新发现了，他发现了这种解包含闭合的时间线，论文发表在*Proc. Roy. Soc. EdinburghA* 57, 135上。范斯托克姆（他的爸爸是梵高的表兄弟）可是个风云人物。他以前是爱丁堡大学的研究生，1939年，他去了普林斯顿高等研究院，希望能成为爱因斯坦的学生。但由于“二战”爆发，范斯托克姆不得不收起他的雄心壮志，投身于反对希特勒的联盟中。他加入了加拿大空军，成为一名轰炸机飞行员。1944年，他又加入了荷兰流亡政府的空军。他是皇家空军轰炸机司令部制定的飞行任务里唯一的荷兰军官，曾多次驾驶哈利法克斯轰炸机去欧洲执行任务。他参加了诺曼底登陆战，任务是轰炸德军的火炮阵地。1944年6月10日，他在一次由400架飞机完成的轰炸任务中，由于座机被防空炮击中而壮烈牺牲，年仅33岁。更多关于斯托克姆的故事，参见：Erwin van Loo, ‘Willem Jacob Van Stockum: A Scientist in Uniform’, June 2004, 网上的英语版本可参见：http://www.lorentz.leidenuniv.nl/history/stockum/VliegendeHollander.html.

[20]采访的文字记录见：http://www.abc.net.au/rn/scienceshow/stories/2006/1807626.htm.

[21] K. Gödel, ‘An Example of a New Type of Cosmological Solution of Einstein’s Field Equations of Gravitation’,*Reviews of Modern Physics* 21, 447 (1949).

[22] M. MacBeath, ‘Who was Dr Who’s Father?’*Synthese* 51, 397–430 (1982); G. Nerlich, ‘Can Time be Finite?’*Pacific Phil. Quart.* 62, 227–239 (1981).

[23]相反，走向未来的时间旅行不成问题，而且经常能在物理实验中见到。这就是相对论中所谓的“孪生子佯谬”。双胞胎中的一个人乘坐高速宇宙飞船旅行，回来以后发现，他比待在家里的孪生兄弟更年轻；实际上，高速飞行也是一场时间旅行，他走进了待在家里的孪生兄弟的未来。

[24]美国哲学家大卫·麦拉曼特（David Malament）对“祖母悖论”的看法是：“认为时间旅行……简直就是荒谬，会导致逻辑矛盾，这种观点的逻辑通常是这样的：如果允许时间旅行，可以让一个人的时间倒流，回到过去的状态，这会导致时空的某些点同时处于P和非P的状态。比如说，我可以回到过去，杀死婴儿时期的我，导致那时的我不可能长大以后变成现在的我。但我想说，这样的反对理由从来不能使我信服……这套逻辑的漏洞在于，它并不能得出它想要得出的结论。显然，如果我回到过去，杀死了婴儿时期的我，就会导致某种矛盾的产生。然而，我们唯一能够从中得出的结论应该是，如果我试着回到过去杀死婴儿时期的我，那么由于某种原因，我肯定办不到。或许在最后一分钟时，我摔了一跤。通常的反对理由并不能说明时间旅行不可能发生，而只是说明如果可以时间旅行的话，也无法实现这样的行动。”参见：Proc. Phil. Science Assoc. 2, 91 (1984)。已故的著名哲学家大卫·路易斯（David Lewis）曾不顾潮流，在“祖母悖论”面前为时间旅行的合理性进行辩护。1976 年，他在一篇评论中说：“时间旅行，在我始终认为是可能的。时间旅行的悖论很奇怪，但不是不可能。这些悖论只能说明，我想没有多少人会怀疑：一个允许时间旅行的世界是个非常奇怪的世界，在许多基本的层面上都与我们所理解的世界格格不入。”参见：‘The Paradoxes of Time *Travel’,Amer. Phil. Quart.* 13, 15 (1976).

[25]对宇宙的旋转施加的最强限制可参见：Roman Juszkiewicz, David Sonoda and J. D. Barrow, ‘Universal Rotation: How Large Can It Be?’,*Mon. Not. Roy. Astr. Soc.* 213, 917 (1985).

第6章
稳态宇宙理论对决大爆炸理论

一旦有人从太空中拍摄到了地球的照片，我们就将在情感上攀上一个新的境界……一旦地球在宇宙中的孤立位置广为人知，那么无论是什么国籍、什么信仰的人，都会获得一个强有力的新观念，它和历史上的任何观念相比都毫不逊色。而且依我看，不远的未来的这种发展可能会带来益处，因为它必然会越来越使国家间的纷争显得徒劳无益。或许，新的宇宙学正是会以这种方式对整个社会的组织结构造成影响……在一位翱翔于太阳系的旅行者看来，地球的全景要比其他行星波澜壮阔得多。

——弗雷德·霍伊尔[1]

亘古不变的宇宙

作为一个科学家，我完全不相信宇宙开始于大爆炸的说法。

——阿瑟·爱丁顿[2]

1948年，人类世界开始从历史上最具破坏力的战争中逐渐恢复。在和平时期，许多科学家重新拾起了从前的工作，而另一些科学家则投入了军事技术的研究，如雷达和核物理，当然不是为了战争，而是纯粹为了科学的目的。战争时期建立起来的亲密友谊和密切合作，往往为日后的科学研究埋下了种子。现代宇宙学中三个响当当的名字就是一例：赫尔曼·邦迪（Hermann Bondi，1919～2005）、弗雷德·霍伊尔（Fred Hoyle，1915～2001）和托马斯·戈尔德（Thomas Gold， 1920～2004）。邦迪和戈尔德都出生在维也纳，经过不同的历程后又都进入剑桥大学继续深造。邦迪在学生时期就表现出了超常的悟性，早在1936年的时候，就有人把他介绍给了爱丁顿，当时这位杰出的科学家正在奥地利首都访问。这个16岁的男孩给爱丁顿留下了深刻的印象，于是爱丁顿就帮他申请去剑桥的三一学院留学。戈尔德并不像邦迪那样从小就野心勃勃，要在学术上大展宏图，而是在父亲的坚持下，才于1937年考上了剑桥，学习工程学。战争爆发以后，由于国籍是敌国的原因，他们俩都被暂时关进了贝里圣埃德的军事基地，然后又在魁北克关了15个月，他们就是在那儿认识的。1942年，他们被放回了英国，戈尔德加入了海军部的雷达研究小组。这个小组由霍伊尔领导，基地设在苏里市的卫特里。在那一阶段的军事研究中，戈尔德做出了一些重大成就，例如他发明了战船登陆的雷达导航系统，在诺曼底登陆时派上了用场，又发现了德国的U型潜艇不用上浮，就可以用一根管子在水下获取空气。

霍伊尔在英国走的是一条传统的教育之路，他从宾格利文法学校毕业后进入剑桥，在那里，这个优秀的毕业生不情愿地成为了一名研究生，而他的导师也同样地不情愿，因为这位导师就是寡言少语的保罗·狄拉克。霍伊尔称他们俩是天生一对，一个是不想要导师的学生，一个是不想要学生的导师。一开始，霍伊尔突击学习了核物理学和量子力学，但后来他的兴趣又转移到了天体物理学的新课题。他相信了朋友瑞

·利特列顿（Ray Lyttleton）的话，天体物理学里充满了令人着迷的未知难题，也缺少像狄拉克那样的天才人物。

邦迪、戈尔德、霍伊尔在结束战时工作后都回到了剑桥。邦迪和霍伊尔都研究天文学，在天文学界已经小有名气。戈尔德研究的是完全不同的东西，他成了一名生物物理学家，研究人类的听觉，并把他在霍伊尔小组研究信号时的经验运用到了新的工作中。他使用了一种特殊的、不借助数学的方法，这让行内的专家大惑不解。尽管这项成果[3]让戈尔德在剑桥大学争取到了一个职位，但在当时的医学界并没有掀起大风大浪。直到20世纪70年代，人们才发现耳蜗内的一种反馈机制确实能产生额外的声音共振①，戈尔德的预言验证了。

① 戈尔德发现，考虑了内耳中的液体产生的阻尼之后，理论上预言的耳蜗共振强度与实际的不符。因此他假设，人体内有一种反馈机制能够抵消这种阻尼的影响。

——译者注

从那时起，邦迪、戈尔德和霍伊尔开始关注宇宙学的进展。邦迪对广义相对论产生了兴趣，并很快发现了不均匀的宇宙模型，这是托尔曼早在20世纪30年代就已经发现了的。但是当三人调研了相关文献，并且在吃饭和喝茶的时候进行了大量讨论以后，他们都觉得勒梅特膨胀演化的宇宙模型不能令人满意，尽管这个模型曾让整个天文学界大为赞叹。勒梅特的宇宙从过去某个时刻开始膨胀，持续不断地演变，并注定会变成一片荒芜的不毛之地，这让他们非常沮丧。

顺着爱因斯坦的思路，他们也假设宇宙从大体上看是均匀的，但他们不相信宇宙的过去和现在有什么不同，也不相信将来会有什么不同。为此，他们提出了一个非常新颖的宇宙模型：宇宙从古至今、从头到尾都是一样的。米尔尼把爱因斯坦关于宇宙的物质均匀分布的假设叫做宇宙学原理。而宇宙从古至今都是一个模样是一个更强的限定，邦迪、古尔德和霍伊尔把它称为完全宇宙学原理。

他们提出的这种原理，并不是说宇宙是静止的，而是说宇宙时刻处于稳态。他们知道，天文学观测表明宇宙在膨胀，因此不可能处于静止状态。但这并不意味着宇宙不可能永远都一样。设想一条大河稳稳地流过，如果你盯着河面上的一点看，就会发现大体上没有什么变化。“宇宙也可以这样吗？”他们这样问自己。

他们希望无论何时，宇宙从平均上看都是一样的。局部上可以存在一些小变化，不断有恒星和行星在形成，然而宇宙所有的总体性质（膨

胀的速率、星系和恒星形成的速率、物质的密度、辐射物质的温度）平均而言，在任何时刻都是不变的。这与膨胀宇宙的标准模型迥然不同。在这些标准模型中，宇宙过去的物质密度和温度都非常高，随着宇宙年龄变大，膨胀的速率也逐渐下降，最终会走向一个冰冷的、没有生命的未来（被误称做“热寂”）或一个灾难性的“大塌缩”。在弗里德曼和勒梅特的宇宙模型中，宇宙的过去和未来有着天壤之别。

为了能够严格地满足完全宇宙学原理，1948年，邦迪、戈尔德和霍伊尔（图6.1）提出了一个激进的新想法：如果要让宇宙膨胀，又要始终保持宇宙的物质密度不变的话，唯一的办法就是要有新的物质不断冒出来，这样才能抵消宇宙膨胀对物质密度的稀释作用。

这个想法听起来就像天方夜谭，但他们辩解道，相比其他宇宙学模型提出的，所有的物质奇迹般地在一瞬间同时冒出来的想法，这已经不算什么稀奇的了。如果新物质真的会不断地产生的话，它们就有可能被物理学家探测到。要保持宇宙密度不变，所需要的新物质生产率非常非常微小——大约是每百亿年从每立方米的空间中产生出一个原子。这个生产率太小了，远远超过任何实验所能直接探测的范围。没有任何一个物理实验室能制造出平均密度像宇宙那么小的真空。然而，尽管你不能指望直接探测到新物质的产生，这个理论也远没有超出天文学观测的检验能力。恰恰相反，这个理论做出的大量言之凿凿的可检验的预测，都被天文学家用观测数据否定了。



图6.1 1960年左右，托马斯·戈尔德（左）、赫尔曼·邦迪和弗雷德·霍伊尔（右）在剑桥的一次会议中

假如宇宙的总体性质真如他们所说，从古至今都没有变过，那么当我们向太空极目远眺（于是会看到很久很久以前发出来的光）时，应该发现宇宙任何时候看起来都一样。不可能存在一个这样的时间段，期间星系开始形成，而在那之前，宇宙中还没有任何星系。20世纪50年代，越来越多的证据推翻了这个重要预言，让人们逐渐从稳态宇宙理论中清醒过来。射电天文学家说，星系发出的无线电波并不是普遍存在于宇宙过去的任何时期。后来，人们发现了一种叫做类星体（全称是“类星射电源”[4]）的新型发光天体，能从像太阳系这么小的区域内发出有如整个银河系大小的能量，这给稳态宇宙模型提出了新的难题。所有类星体似乎都与我们保持特定的距离，这说明它们形成于宇宙历史上一个短暂的特定的时期。稳态宇宙模型无法解释这个现象，而在弗里德曼和勒梅特的演化宇宙中，这个现象非常自然，这表明当宇宙膨胀到一定时期，各种条件成熟了以后，星系、射电源和类星体这样的天体就开始形成了。

当时没有人注意到，这其中有一个奇怪的巧合。在弗里德曼–勒梅特宇宙模型中，宇宙的膨胀速度大约是宇宙年龄的倒数。在稳态宇宙模型中，膨胀速率和宇宙年龄没有任何关系，因为宇宙的年龄无穷大——

宇宙一直都存在。稳态宇宙的膨胀速率可以任意取值。实际上这无关紧要。因此，宇宙实际膨胀速率的倒数，跟太阳这样的典型恒星的年龄相近，在稳态理论中完全是巧合。然而，在弗里德曼和勒梅特的演化宇宙模型中，这件事完全是自然而然的。先有恒星，才能有天文学家。于是我们可以设想，当我们观测宇宙的时候，宇宙的年龄已经足够大，足以让恒星安顿下来，并把燃烧氢元素的核反应也稳定下来（需要大约一百亿年）。这样，我们就会发现，宇宙的膨胀速率接近于恒星年龄的倒数。①

① 哈勃常数表示宇宙膨胀速度的大小，量纲是速度（千米每秒）除以距离（百万秒差距），因此哈勃常数倒数的单位是时间。在许多演化宇宙的模型中，哈勃常数的大小往往和逝去的时间呈反比，因此其倒数接近宇宙的年龄和恒星的年龄是非常自然的结果。——译者注

邦迪和戈尔德在1948年提出的稳态宇宙论中没有多少方程。[5]这个理论要满足对称性和时空的均匀性原则。他们坚持认为，宇宙从古至今、从头到尾都应该看起来差不多是一个模样。这通常被看成是一种对哥白尼原理的扩充，它不允许我们认为自己正处在宇宙的一个特殊位置。在稳态宇宙论看来，这个位置指的是时空中的位置，而不是仅仅代表空间位置。

这个原理的限定性非常强，只有四种类型的宇宙能够满足其要求。第一种是什么都没有的静态宇宙——没有物质，没有辐射，没有引力，只有时间和空间。这太没意思了。第二个是爱因斯坦最开始提出的静态宇宙。这个宇宙不膨胀，于是所有地方一直都是一样的。现实并非如此，所以邦迪和戈尔德没有选择这个模型。生命的出现要求宇宙得发生一些变化。根据热力学第二定律的预言，宇宙总体的混乱度和复杂度也要不断增加。第三个选择又是一个不膨胀的宇宙，就是哥德尔发现的那个奇特的旋转宇宙。这个宇宙在匀速转动，无论你在何处，看起来都一样。这个旋转宇宙明显不是我们所生活的宇宙。最后一个宇宙终于满足了邦迪和戈尔德的要求。它就是最先由德希特发现的、空间曲率为零的宇宙。不像其他三个，这个宇宙会膨胀，而且膨胀速率从古至今、从头到尾都是一样的。活在这个宇宙的天文学家，无法通过观测来确定现在是“何时”。[6]这个宇宙没有开端，也没有终点。

邦迪和戈尔德提出稳态宇宙模型不久，霍伊尔也发表了一个自己的版本。[7]他引入了一种物质均匀产生的机制，可以抵消宇宙膨胀的稀释作用。这个机制会直接导致与德希特宇宙类似的演化行为。

这种模型有一些有趣的特点。首先，它解释了为什么宇宙是均匀

的、膨胀速率是各向同性的。如果你向稳态宇宙中引入任何不对称性，其产生的影响都会迅速减弱，宇宙的膨胀也会迅速恢复到各向同性、均匀的状态中来。在微扰的影响下，稳态宇宙是稳定的。[8]打个比方，假如你暂时把铅笔倒立起来，你会发现这个姿态非常不稳定，任何风吹草动都能使它摔倒。但如果你把一个玻璃弹子放在碗底，并把它推离平衡位置的话，你就会发现它在平衡位置周围飞快振荡，最后回到原点。这就是稳定的物理系统。

稳态宇宙模型引发了巨大的争议，因为它第一次抛弃了爱因斯坦的引力理论，试图仅仅用对称性原则描述宇宙。邦迪、戈尔德和霍伊尔似乎对此不以为意，因为他们的理论能够做出许多简单的预言，很容易被天文学观测检验。霍伊尔提出的物质不断产生的现象，在广义相对论中并不存在，是一种新的物理机制。随后在1951年，英国天文学家威廉·麦克里证明，爱因斯坦的理论中可以很容易地引入霍伊尔的“创生场”，而不必改变其本质。这不过是用另一种形式构造爱因斯坦早先提出的宇宙学常数。宇宙学常数的大小和新物质产生的速率相等。这个速率始终保持不变，所以说它的行为很像德希特宇宙，如图3.4所示。

整个20世纪50年代，稳态宇宙模型一直被看作是传统的“大爆炸”理论有力的竞争者。“大爆炸”一词还是霍伊尔1949年一次在做客BBC第三套广播讲宇宙学时提出的呢！当时观测天文学的前沿是射电天文学，打头阵的是剑桥的天文望远镜，由马丁·赖尔（Martin Ryle，1918～1984）负责实施。赖尔在观测中应用了由他本人发展的新技术，可以提高天文观测的质量。尽管霍伊尔和赖尔在战争期间都曾研究过雷达技术，可他俩的关系并不好，互相怀疑对方故意找茬，想推翻自己的研究成果。例如，赖尔试图用观测数据推翻稳态宇宙的预言，而霍伊尔则质疑宇宙射电源观测结果的精确性和赖尔做出的解读。最终，赖尔的数据几乎征服了所有人，大家都开始相信天然射电源并不是从古至今都普遍存在的，而这违背了稳态宇宙模型的要求。下面这首诗是由乔治·伽莫夫的妻子芭芭拉写的，从旁观者的角度记录了这场争论：

《赖尔辩霍伊尔》

赖尔辩于霍君：
“尔等夙兴夜寐，
我谓宝山空回。
磐石虽稳，
难阻东流之水。
焉能颠倒是非。
我有巡天远镜，
离相贪嗔寂灭，
化解痴心无形。
劝君一言：
既知天高海阔，

自有云淡风轻。”
霍君反唇相讥：
“勒梅特伽莫夫，
陈年旧事休提。
肉食者鄙，
秕言开天辟地，
焉长他人志气？
此理昭然易见，
绝无海枯石烂，
亦非盘古开天。
我辈三人，
绵绵不绝此念，
精疲力竭亦然。”
“非也非也非也，”
赖尔目眦尽裂，
怫然强忍不悦，
“芸芸星系，
放眼穷极视界，
无不森然布列。”
“竖子不足为教！”
霍君如雷暴跳，
词穷理屈又表，
“生生不息，
无论暮暮朝朝，
恒久长存其妙。”
“任尔唇剑舌枪，
我有铁壁铜墙，
切莫贻笑大方，”
赖尔再道，
“不出一时半晌，
必来缴械投降！”[9]

赖尔利用剑桥的射电望远镜获得了大量天文观测证据，邦迪、古尔德和霍伊尔三人却无动于衷，仍在射电天文学家的咄咄攻势中做困兽犹斗。这场争论在宇宙学历史中被称作“大爆炸对决恒稳态”，尽管实际上这主要是英国天文学界的事，没怎么影响到美国。[10]当然，这件事也体现出，英国和美国宇宙学家的研究方法存在一个有趣的差异。英国的宇宙学家看重科学方法论以及其中蕴涵的科学哲学原理，把它们用来估量一些概念的地位和隐含意义，例如邦迪、古尔德和霍伊尔的完全宇宙学原理。相比之下，奉行实用主义的美国式研究则避开了这些争论。[11]由于霍伊尔在广播节目中发挥出色，而且他的书也十分畅销，在英国这场争论简直是家喻户晓，直到今天，仍然有非专业人士以为稳态宇宙模型和大爆炸理论是棋逢对手。

20世纪 80年代，我曾和霍伊尔有过一段时间的接触。当时我们在意大利博洛尼亚参加一个现代宇宙学历史的研讨会，被安排住在一起。我们讨论了许多那段稳态宇宙模型的命运还不甚明朗的时期，以及在所

有人都放弃了这个理论之后的二十多年间，他是如何看待这些否定性证据的。最让他困扰的是，这个物质不断产生的过程，应该冒出等量的物质和反物质。在他看来，物质和反物质的平等性是一个大问题[12]，因为当时没有任何证据表明存在反原子、反行星或反恒星：宇宙中的物质具有压倒性的优势。的确，我们能在实验中制造出反粒子，能在宇宙射线中探测到正电子，但宇宙主要还是由物质构成的，至少在我们周围的地方是如此。

非常值得注意的是，1948～1952年间，由于最遥远退行星系的距离被科学家算错了，所有的膨胀宇宙模型都被这个数据误导了，于是宇宙的年龄也被大大低估了。在这段时期内，这个数据继而导致了一个悖论，膨胀宇宙模型预言的宇宙年龄比其中最古老恒星的年龄还要略微小一些。就是在这种混乱的情况下，稳态宇宙模型诞生了。于是，它的对手大爆炸理论处境艰难。大爆炸理论不太能描述宇宙的细致特征，因此无法成为一个比稳态宇宙模型更好的理论。爱因斯坦对这个艰难处境的反应值得玩味——而且现在回过头来看，他的反应完全正确。面对恒星年龄比宇宙年龄还大的尴尬，他断定更可能是恒星演化理论错了："在我看来，无论如何，恒星演化理论所依赖的基础不如引力场方程牢固。"[13]

1952年，由于沃尔特·巴德（Walter Baade，1893～1960）的出场，境况开始变化。巴德是一个德国天文学家，从1931年起在美国生活和工作。他怀疑当时造父变星的绝对亮度有误，于是用坐落在帕洛玛山的200 英寸望远镜来验证自己的怀疑。天文学家用造父变星来测量河外星系的距离，进而根据哈勃–勒梅特定律，描述遥远天体的退行速度。当时的观测结果表明，虽然仙女座星系看起来和银河系如此相像，但它的尺寸远比银河系小得多，于是巴德就产生了怀疑。巴德嗅到了蛛丝马迹，并通过一丝不苟的工作，最终发现了问题所在。他证明了造父变星比人们想象的更亮。因此，天文学家们大大低估了遥远恒星和星系的距离。巴德证明，天体的距离都应该翻倍，而宇宙的年龄也应该是原先预计的两倍——36亿年，而不是 18亿年。到了 1954年，关于星系的距离有了进一步修正，于是宇宙的年龄增加到了54亿年。最近，借助于测量遥远天体距离的新方法，以及新的地面和太空望远镜的出色观测能力（如哈勃望远镜和斯皮策望远镜），宇宙的年龄终于被确定了下来。如今，对膨胀宇宙年龄的最佳估计是 137 亿年，不确定度只有1亿年。

桌面上的宇宙

想当年，我们的科研预算通常都很紧张。我们的手头自然也就非常拮据。例如写1941年的那篇论文时，我不得不一个人挑起所有重担，包括重建实验室，安装所有电子仪器。这很花时间，不过当发现自己什么都会做时，你就会非常惬意和满足。现如今，科学家都被宠坏了，尤其是在发达国家。

——埃里克·霍尔姆伯格[14]

在当今社会，电子计算机非常普及，而且容易上手，以至于很难想象没有它们，世界会变成什么样子。天文学家们利用大型计算机来模拟一系列复杂的物理过程，这能够描述恒星和星系如何形成，以及是如何在天上分布的。1941年，埃里克·霍尔姆伯格（Erik Holmberg，1908～2000）还是瑞典隆德皇家大学天文台一名年轻的天文学家，是他第一次模拟了宇宙。他并没有使用现代意义上的计算机（当时还没发明），实际上，他自己造了一台模拟器，可以模仿引力的行为，然后进行观察和测量，看看多恒星系统在引力的作用下会如何演化。

我们所生活的三维世界有一个惊人的特点，那就是许多自然的基本作用力和效应，会随着彼此距离的增大而以平方反比的规律迅速衰减。[15]引力、磁力、静电力和光照的强度都符合这个规律。霍尔姆伯格想知道，两个被恒星环绕的星系相互靠近时，会发生什么情况。来自另一个星系的引力是如何改变这个星系中的恒星轨道，继而改变星系形状的呢？由于宇宙在膨胀，从前的星系要比现在靠得更近，因此狭路相逢的可能性也比今天高得多。[16]

霍尔姆伯格决定利用光照的强度和引力一样都遵循平方反比定律的特点。他用一组灯泡来模拟多体系统中的引力分布。在一个暗室的桌面上，他安装了两个扁平的星系模型（就像银河系），每个星系里有37个灯泡，分别排列在各自的同心圆结构上。每一个灯泡的亮度都代表了对应恒星的质量，而且根据距离圆心的远近，这些灯泡的亮度都要做相应的调整，以符合恒星分布的实际情况。每个灯泡的四个面上都有一个传感器，每个面都会所接收到不同距离外的灯泡照射，然后他用光电池和电流计来测量照射的强度。霍尔姆伯格从当地的工厂定制了这些特殊的灯泡。两个灯泡组成的“星系”各自旋转着，同时沿着直线彼此靠近。这样，霍尔姆伯格通过测出照着在每一个灯泡上的光线强度，并计算出总体影响（“合力”的大小和方向），就能知道恒星下一步会如何运动。所有74个灯泡都移动到新的位置，在新的位置上受到的光照强度又决定了灯泡下一步的运动。利用这个精巧的装置，霍尔姆伯格就能够建立一个模型，来模拟两个恒星系统对彼此的引力相互作用，以及一步一步的演化过程——这个过程过于复杂，根本无法用铅笔、草稿纸和牛顿定律来求得精确结果。[17]

图6.2和图6.3摘自霍尔姆伯格的论文。[18]他发现两个星系狭路相逢的时候，彼此之间的引力最强，产生的形变也就最严重，而且经常会产生臂状结构，让人不禁想到许多扁平星系都有旋臂，例如银河系。当两个星系互相做逆时针转动时，这个臂状结构就朝外弯曲，指向旋转的对称轴，但如果是顺时针的转动，臂状结构就指向相反的方向。他还断定，

中小型星系经常会被大型星系的强大引力所俘获，彼此混合在一起，组成了一个新的大型星系，形状变成椭球形而不是螺旋型。

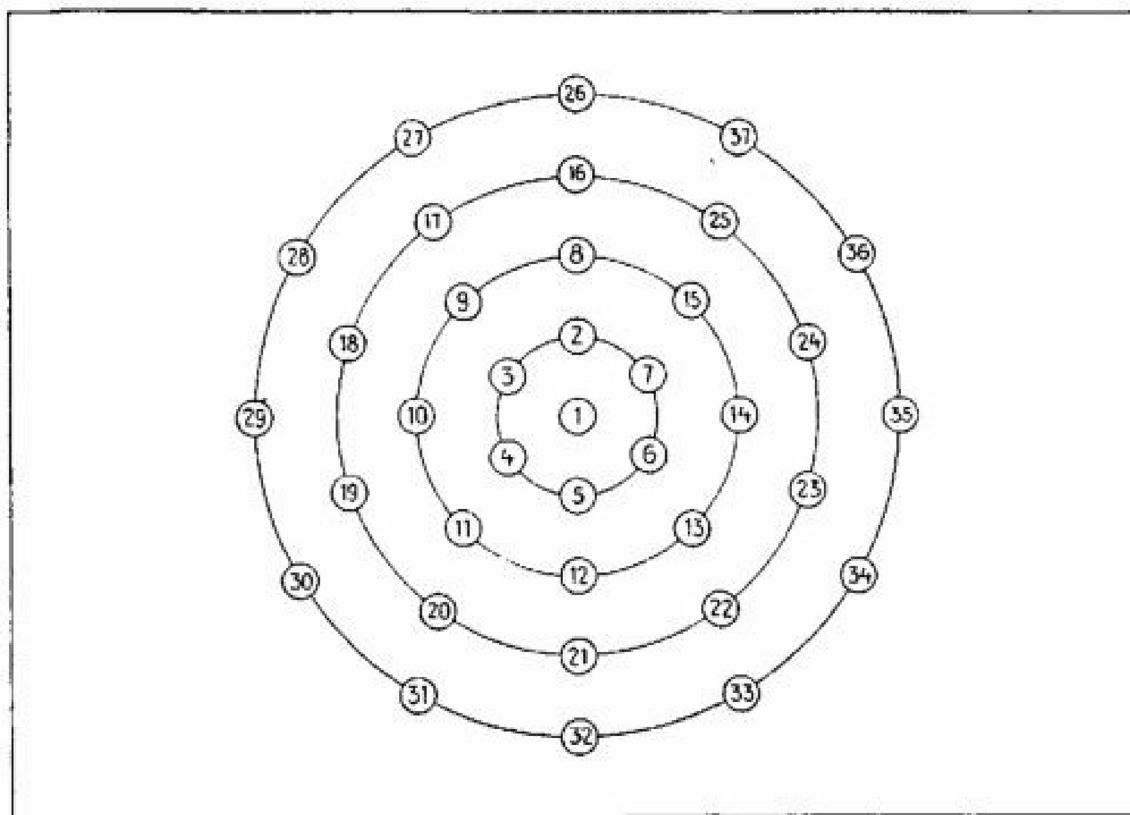


图6.2 埃里克·霍尔姆伯格用电灯泡模拟星系的演化。在他的这幅图中，灯泡被安置在直径80厘米的同心圆上，一共37个灯泡。他做了两个这样的模型，每个模型都代表一个由恒星组成的星系。调节灯泡两端的电压就能改变每个模型中灯泡“质量”的分布（即亮度）。灯泡的直径是8毫米，比灯泡之间的距离10～20厘米小得多

霍尔姆伯格富有想象力的工作给宇宙学研究中引入了一个新的工具。直到20世纪70年代，这种工具才再次出现，并伴随着计算机技术的革命，规模越来越大，模拟能力越来越强。借助这些新工具，人们可以建立宇宙的计算机模型，或者是一些经过简化的模型。将简单的引力定律编成程序，就能计算膨胀宇宙模型中任意数量的质点之间的引力，这样计算机就能跟踪这些质点的演化，了解它们如何结团，弄清楚星系初次形成的时期，或者它们何时会与邻近星系碰撞。最终，我们希望能够了解恒星在星系中如何形成的所有麻烦的细节。

图6.3 两个埃里克·霍尔姆伯格的电灯“星系”模型，每个模型都顺时针转

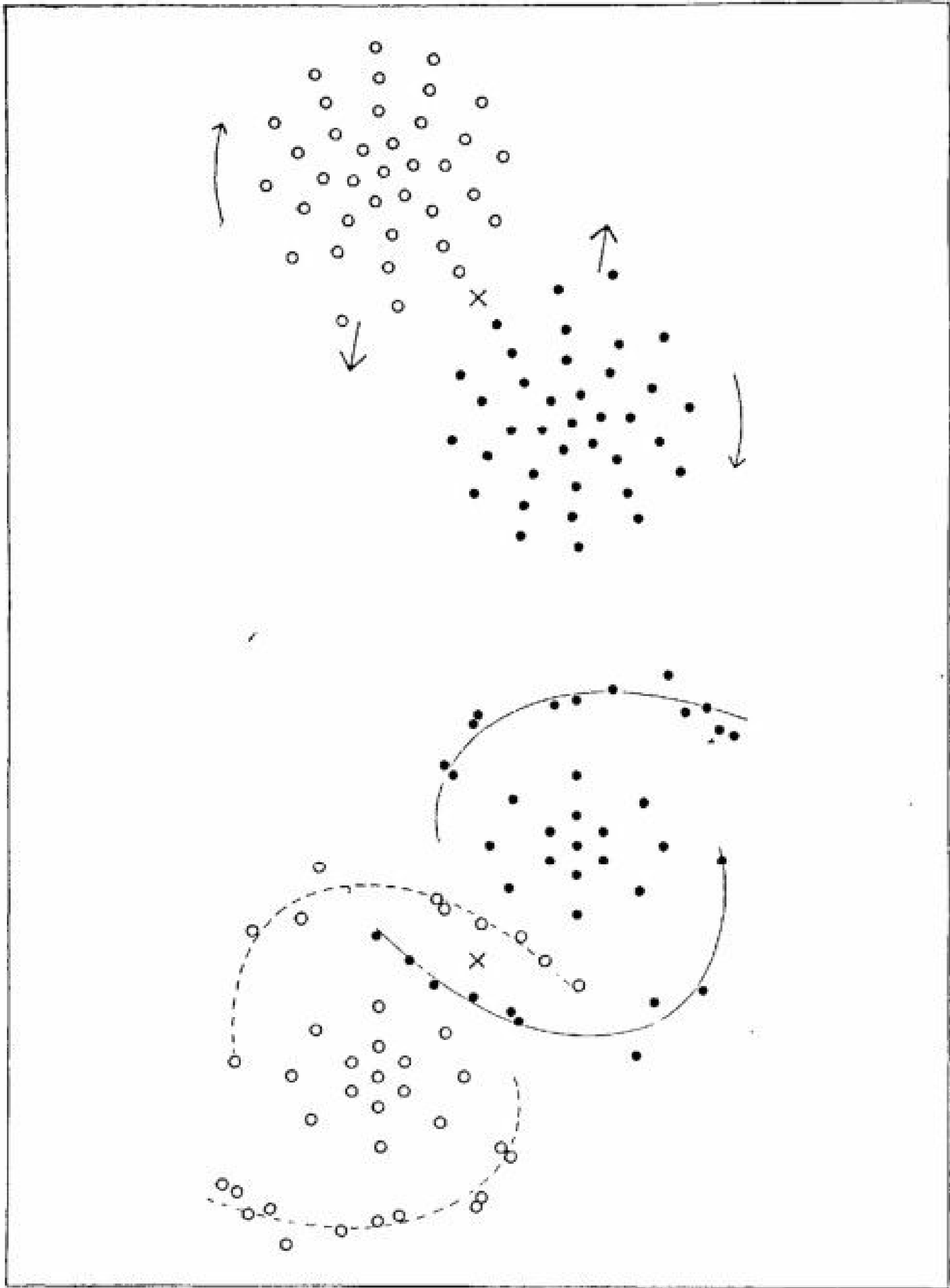


图6.3 两个埃里克·霍尔姆伯格的电灯“星系”模型，每个模型都顺时针转

动，之间的距离等于星系的直径。星系的引力相互作用导致了旋臂的产生，并且指向星系转动的方向。短箭头表示两个星系靠近最近点时的运动方向，弯曲的箭头表示转动方向

今天，在全球最强大的计算机上，宇宙学家们可以模拟不同类型的宇宙，从中预言我们应该在望远镜里观测到什么现象，或是解读一些观测数据的统计含义。鉴于世界上曾经存在两类天文学家，一类是观测天文学家，一类是理论天文学家，现在可以说有了第三种。计算天文学家的能力与众不同，他们能编写错综复杂的计算机代码，生成精彩绝伦的图片或视频，从而展现不同的宇宙模型的不同特点。[19]理论天文家作出预言，计算天文学家检验预言，观测天文学家则负责用望远镜来寻找检验的结果。

带电的宇宙

……由于时常需要把这儿的灯泡换换，又把那儿的晶体收音机调调，我就学到了扎实的电学实践知识。但这些知识并不全面——比如，我还是不明白为什么不能用一个电吉他煮鸡蛋。

——凯斯·沃特豪斯（1929～2009，英国作家）[20]

在发展稳态宇宙理论的同时，1948年，赫尔曼·邦迪还提出了一个非常不同的宇宙模型。他设想，这个模型可能有稳态的形式，但这种形式对宇宙膨胀来说，既非充分也非必要。他和剑桥的同事瑞·利特列顿共同发展了这个想法，并注意到两个像质子这样的带电粒子，产生的静电斥力要比引力大得多（实际上大10^{39}倍）。

通常，人们假设氢原子呈电中性。氢原子包含一个质子和一个电子，两者的电荷大小相等，符号相反：质子的电荷是+e，而电子的电荷是Ήe，所以总电荷是零，原子呈电中性。但是邦迪和利特列顿假设，质子和中子的电荷之间有极其细微的差别，只不过我们还没发现。这样的话，氢原子都会带有一个总的电荷，每个原子的总电荷都符号相同。如果静电的排斥力大于氢原子之间的万有引力的话，它们就会相互排斥（图6.4）。

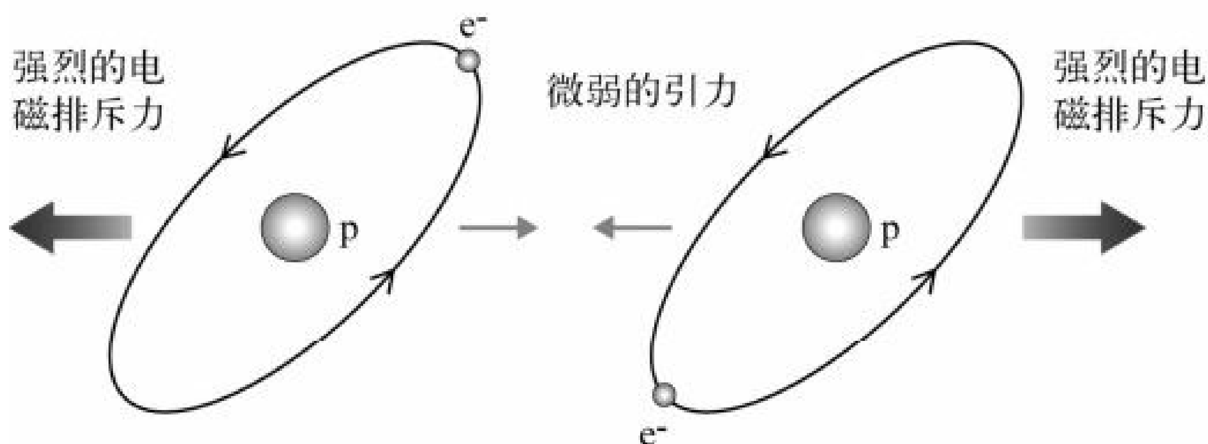


图 6.4 两个氢原子之间有着微弱的引力。但如果质子（p）的电荷和围绕它的电子的电荷不相等，那么每个原子都会带上一个净电荷。每个原子都带同种电荷，于是相互之间就会产生了强烈的电磁排斥力

既然静电的排斥力远大于引力，他们认为，只要质子的电荷比电子电荷的绝对值 e 大 10^{-18}的话，原子之间的排斥力就足以解释宇宙的膨胀。

当这个新奇的观点[21]被提出时，当时普遍接受的对电荷差异的实验约束比邦迪和利特列顿提出的大上 100倍（小于 10^{-16}）。但渐渐地，更精确的实验结果显示，正负电荷之间的差异小于e的10^{-20}。起初，邦迪和利特列顿拒绝接受这些结果，因为他们坚持认为，在原子尺度上，电荷差异的效应非常微弱，而且实验中用到的电磁场对结果本身会有什么影响也没有被充分考虑在内。但渐渐地，实验证据消除了这些担心，这个理论也就没法继续维持了。[22]最后，1960 年，美国科学家约翰·戈登·金研究了两个氢原子组成的氢分子的电荷，发现实验结果所允许的电荷差异，要比解释宇宙膨胀所需要的差异至少小40倍。[23]

这个理论从来没有成为宇宙学研究的主流，但这是一个很好的例子，体现了亚原子物理学是如何进入爱因斯坦的相对论所把持的引力领域，以及它是如何改变我们对宇宙膨胀的认识的。

在大西洋的另一侧，有一个宇宙学的研究小组提出了另一个宇宙模型，他们更关心在爱因斯坦方程组的描述下，膨胀宇宙的早期阶段发生了什么样的物理过程。

炙热的宇宙

《热情似火》

——比利·怀尔德执导，伊蔡克·戴蒙德编剧[24]

在英国，稳态宇宙模型吸引了如此多的公众注意力，以至于它的对手几乎淡出了公众的视野。然而，1948年以后，对大爆炸模型的关注逐

渐多了起来，这种模型认为过去的宇宙要比现在的状态更热更致密。爱因斯坦的理论预言了不同类型的膨胀宇宙模型，天文学家们已经在研究如何区分这些模型，而物理学家们则开始研究早期宇宙不同寻常的物理环境。勒梅特率先考虑了胚胎宇宙的物理状态，而苏联流亡者乔治·伽莫夫和他的学生真正把这个想法落到了实处，他们对相关问题的证明成了对宇宙过去最实用而可靠的指南。

伽莫夫生活在一个风云变幻的时代，而他的人生也跌宕起伏、多姿多彩。他成长在俄国革命风暴席卷下的奥德萨，然后进入了当时的圣彼得堡国立大学，并且在亚历山大·弗里德曼的课上学习了广义相对论。才二十出头，他就在核物理领域作出了重要贡献。伽莫夫性格外向，从不循规蹈矩。他好像有一种过人的能力，能够结识三教九流各色名人，从莫洛托夫、布哈林、托洛斯基到爱因斯坦、玻尔、弗朗西斯·克里克，而且通常是在重要的场合。[25]

伽莫夫逃离他的故乡奥德萨、流亡到美国的故事可谓惊世骇俗。由于当局对知识界的干预日趋严重，伽莫夫深感恐惧，在1932年决定和第一任妻子柔儿（Rho）穿越守备松懈的边境，逃离自己的祖国。他们从克里米亚半岛的南端出逃，希望走私贩子能帮他们渡过黑海，到达 270 公里之外的土耳其。在这个逃亡计划中，他们需要一艘可折叠的小橡胶艇，以便漂洋过海。他们打算把橡胶艇和桨分别放在两个帆布包里，背在身上。尽管当时的苏联食物短缺，他们还是准备了一周的食物，并临时带上了一些漂浮用具（足球胆）、一个水泵和一个指南针。他们进入土耳其的时候不计划带身份证，而是声称自己是丹麦人，因为伽莫夫有一个丹麦的旧驾照，这样就可以想办法进入伊斯坦布尔的丹麦大使馆。1932年初夏，他们在克里米亚海岸的一个科学院度假村订了房间，把小艇也带去了，进行了渡海演习。

本来，他们打算等到满月的时候再出发，这样对逃跑有利。可当时海上风平浪静，他们便打算立刻出发。头两天，海上之旅一切顺利，不久风浪渐渐大了起来，海水不断灌进小艇，柔儿不得不让水泵昼夜工作，以防小艇沉入大海。最后，船还是翻了，他们发现自己离海岸不太远，面前还能看到一些渔夫。非常遗憾，这些渔夫不是土耳其人：这对夫妇离他们出发的地方仅有 110公里。在渔夫们的帮助下，他们又回到了原点。幸运的是，没有人怀疑他们编的瞎话：他们是赛艇爱好者，由于经验不足，加上天气恶劣，不小心被冲到了很远的地方。真相实在是太疯狂了，没人会相信。

伽莫夫夫妇并没有放弃逃离的念头。两年以后，在尼尔斯·玻尔的邀请下，他们赴比利时参加了一个物理学会议，会后并没有返回苏联。乔治得到了密歇根大学安娜堡分校的一个职位，并在美国度过了余

生。[26]

得益于对核物理学的深刻了解，伽莫夫设想，宇宙刚开始膨胀的那几分钟里，也许就像一次巨型的核物理实验。经过了一些错误的尝试之后，在伽莫夫的两个年轻研究生拉尔夫·阿尔珀（Ralph Alpher，1921～2007）和罗伯特·赫尔曼（Robert Hermann，1914～1997）的正确努力下，这个理论的核心思想初具雏形。

1948年夏天，伽莫夫证明，宇宙的膨胀进行了大约一百秒之时，质子和中子就开始相互结合，形成氘核。[27]六个月以后，阿尔珀和赫尔曼发展了伽莫夫的想法，在一篇写给《自然》杂志的信中，他们计算了宇宙的温度和密度随时间变化的规律，这样就能将任何一个历史时期的宇宙温度和密度同今天的数值相联系。[28]

阿尔珀和赫尔曼证明，如果宇宙膨胀的过程是各向同性而均匀的，那么从起初的炙热状态开始，用物质密度除以辐射温度的立方，得到的比值始终是个常数。这意味着他们可以确定当宇宙刚刚诞生两分钟（其辐射温度高达十亿度）时这个比值应该是多少，以避免当时的核反应产生过多的氦元素而不符合我们今天观测到的数值。这样知道这个比值的大小，又知道现在宇宙平均密度的观测值后，他们就能推导出现在的辐射温度大小。他们估计这个数值应该是5开尔文①左右。这是历史上最重要的科学预言之一。这样的话，天文学家就能够通过观测来验证，宇宙是否如大爆炸理论所说，曾经处于极度炙热而致密的状态。这些计算最后证明，如今的宇宙中应该存在高温辐射遗留下的痕迹。

① 开尔文是国际单位制中的温度单位，以下简称开。5 开相当于零下 268.15 摄氏度。——译者注

唉，可惜当时好像没人关注这几篇论文。1948年的稳态宇宙之争吸引了人们的注意力，而且欧洲的科学家从没听说过阿尔珀和赫尔曼。美国的情况也好不到哪儿去。20世纪60年代中期，普林斯顿的一位顶级物理学家罗伯特·迪克，打算开展一个探测宇宙辐射的项目，可他竟然没有读过伽莫夫、阿尔珀和赫尔曼在世界顶级期刊上发表的那些文章，更不知道他们已经就相关问题作出过预言。

1965年，阿诺·彭齐亚斯（Arno Penzias）和罗伯特·威尔逊（RobertWilson），两个顶级无线电工程师，正在新泽西州霍姆德尔为贝尔实验室校准一个接收天线，用以追踪“回声”通讯卫星。最终，他们发现了来自宇宙大爆炸的辐射痕迹，尽管他们当时并不清楚自己探测到的东西是什么，也不知道以阿尔珀、赫尔曼和伽莫夫为代表的宇宙学家已经预言了这种辐射的存在。他们探测到的是波长为7.35厘米的无线电

噪声，这相当于温度为3.5开±1.0开的辐射。迪克和他的合作者当时正在从理论上估算辐射的温度，于是很快知道了此事。迪克以前的一个学生，詹姆斯·皮博（James Peebles）估算了这个温度的大小约为 10开，并写成了论文向学术期刊投稿。这篇论文最终被期刊选择的审稿人（正好是阿尔珀）拒绝了，因为文章并没有原创思想，其中的结果也早就被阿尔珀和

赫尔曼算出来了。然而，迪克的研究团队好像没有理会这个信息，又在《天体物理学报》彭齐亚斯和威尔逊所发表的低调的公告旁边发表了一篇论文，专门解释这次探测到的辐射。关于这次集体失忆导致科学研究进度落后了许多年的经历，科学史学家写下了不少的论述。1975年，迪克在自传性笔记中写道：

> 关于火球的热辐射理论，我们的工作存在不幸和令人尴尬的一面。我们没有做足文献搜索的功夫，而与伽莫夫、阿尔珀和赫尔曼的重要论文失之交臂。对此我应该负主要责任，因为团队中的其他成员太年轻，不可能了解这些老文章。早些时候，我在普林斯顿听过伽莫夫的讲座，但我记得他的宇宙模型是寒冷的，开始时只有一些中子。[29]

阿尔珀和赫尔曼在余下的时光中，仍然对所受到的不公平待遇耿耿于怀。后来他们去了工业界工作，对宇宙学后来的发展影响不多。[30] 1955年，阿尔珀离开约翰·霍普金斯大学，去了通用电气公司工作，并于2007年去世。1956年，赫尔曼开始领导通用汽车公司研究实验室的物理部，然后又去了德克萨斯大学当教授，并在那儿做出了关于交通流量的数学理论的一流研究工作。后来他又成了一位相当有造诣的雕刻家，制作木质微缩模型，在美国多处都有展览。他于1997年去世。

Sept 29th 1963

Gamow Dacha
785 - 6th Street
Boulder, Colorado

Dear Dr. Penzias,

~~Send~~ Thank you for sending me your paper on 3 °K radiation. It is very nicely written except that "early history" is not "quite complete". The theory of, what is now knows as "primeval fireball" was first developed by me in 1946 (Phys Rev. 70, 572, 1946; 74, 505, 1948; Nature 162, 680, 1948). The prediction of the numerical value of the present (residual) temperature could ~~can~~ be found in Alpher & Herman's paper (Phys. Rev. 75, 1093, 1949) who estimate it as 5 °K, and in my paper (Kong Dansk. Ved. Sels. 27 no 10, 1953) with the estimate of 7 °K. Even in my popular book "Creation of Universe" (Viking 1952) you can find (on p. 42) the formula $T = 1.5 \cdot 10^{10} / t^{1/2}$ °K, and the upper limit of 50 °K. Thus, you see the world did not start with almighty Dicke.

Sincerely G. Gamow?

图6.5 乔治·伽莫夫在1965年（日期写成了1963年）写给阿诺·彭齐亚斯的信中指出，他在彭齐亚斯和威尔逊发现微波背景辐射以及罗伯特·迪克和他的团队给出实验解释之前就作出过相关预言[31]

尽管被遗忘了很长一段时间，但由于在微波背景辐射和大爆炸理论方面作出的杰出贡献，最终阿尔珀和赫尔曼还是获得了许多奖项，受到了大量赞誉。2005年，在去世两年前，阿尔珀获得了美国国家科学奖。获奖说明这样写道：

GENESIS OF THE BIG BANG



RALPH A. ALPHER

ROBERT HERMAN

图6.6 罗伯特·赫尔曼（左）、乔治·伽莫夫和拉尔夫·阿尔珀（右）。伽莫夫从一瓶形成宇宙的原初物质（ylem）中冒了出来，不过这看上去更像一瓶酒

为了表彰他在核合成领域的突破性工作、他关于宇宙膨胀会留下微波背景辐射的预言，以及他建立的大爆炸理论模型。

彭齐亚斯和威尔逊发现微波背景辐射的事件是一个历史性转折点，从此我们得以更加深入地理解宇宙，同时也对用爱因斯坦方程组描述宇宙行为的方法更加有信心。弗里德曼和勒梅特的最简单的膨胀宇宙模型告诉我们，宇宙每时每刻的温度应该是多少。知道了这个简单的事实之后，物理学家就能搞清楚从宇宙诞生几秒钟起到现在发生的一连串事件。他们无法搞清楚所有发生的事情（远没有做到这点），但很快他们就建立了早期宇宙的行为模型，描述了温度和密度是如何降低的、核反应过程是何时发生的，以及辐射冷却到什么时候原子和分子才开始形成的。

宇宙微波背景辐射的发现也为稳态宇宙理论敲响了丧钟。突然之间，认为射电源不会存在于宇宙每一个历史时期的论调大获全胜，因为他们打出了微波背景辐射这张王牌。在大爆炸理论中，这个远古的现象非常自然（这也是大爆炸理论的预言之一），而稳态宇宙理论就无法解释。以霍伊尔为首的稳态宇宙理论的支持者，继续为了这个理论负隅顽抗。他们试图证明，新发现的微波背景辐射只不过是不久之前在银河系中形成的。然而，这些英勇的抵抗者最终全军覆没了：相比宇宙中的物质密度，微波背景辐射中光子太多了，不可能来自任何已知的辐射源。1967年，科学家们发现，从天空中的任何方向测到的微波背景辐射强度都相同，不同方向之间的强度差别不超过千分之一。不可能通过几个局域射线源的叠加来产生这种平滑性，除非射线源的数量巨多无比，可如果是那样的话，我们早就该探测到这些射线源了。

奇怪的是，霍伊尔本人也在热大爆炸宇宙学中扮演了关键角色。利用核物理学方面丰富的专业知识，霍伊尔走得比伽莫夫、阿尔珀和赫尔曼更远。他预言了宇宙中最轻的几个化学元素的丰度，这些元素应该在大爆炸后的最初几分钟内被制造出来。他先是和剑桥的罗杰·泰勒（Roger Taylor，1929～1997）于 1964年发表了一篇开创性的论文[32]，接着在1967年又和加州理工学院的威利·福勒（Willie Fowler，1911～1995）及其学生罗伯特·瓦格纳（Robert Wagoner）合作发表了文章，在这些文章中，霍伊尔预言了不同的大爆炸宇宙模型中，轻元素（特别是氘、氦的同位素以及锂）的完整丰度分布图。

1953年，日本天文学家林忠四郎（1920～2010）一个关键的小发

现，为这些令人印象深刻的预言补上了缺失的一环，这也是他在恒星演化理论当中最著名的工作。每个物理学家，例如伽莫夫，在尝试预言大爆炸产生的化学元素丰度时，都遇到了一个重要的问题：在宇宙膨胀的最初时刻，质子数和中子数的比例是多大？任何原子核都是由这两种粒子组合而成（例如，氦-4的原子核是由两个质子和两个中子组成的），所以化学元素最终的丰度当然是由质子和中子的初始比例决定的。如果一开始全部都是质子，那么最后你得到的产物100%都是氢原子，但如何一开始不是这样，而你又不知道质子中子的比例的话，你能预测最后的结果吗？

林忠四郎指出，如果考虑到自然界中存在一种引起放射性元素衰变的弱相互作用，这个问题就迎刃而解了。当宇宙还是一个出生不到一秒钟的婴儿时，其中的温度高达一百亿度以上，此时这种弱相互作用会导致质子和中子的数量相等，形成一种平衡状态。只要处于完全平衡的状态中，质子和中子的比例就只由温度的高低决定。因此，你不需要知道宇宙"开始"的时候发生过什么，也不需要知道其他不确定的细节，就能计算质子和中子的比例。事实上，中子的质量比质子大了一点点。当宇宙诞生一秒钟左右，温度下降到一百亿度附近时，由于制造一个中子需要的能量多了一点点，一开始的时候，质子的数量也就比中子多了一点点。然而，这其中的数量差异并没有多少机会发展壮大，因为宇宙在不断地膨胀，其中起主导作用的弱相互作用反应速率太慢①，最后反应停止了。此时宇宙中的中子数和质子数就固定下来，约为1∶6。

① 俞允强的《物理宇宙学讲义》提到，弱过程的发生率是$\Gamma_{弱} \propto T^5$，宇宙膨胀速度是$H \propto T^{3/2}$。因此温度下降时，弱过程的发生率降得更快。当两者大小相等时，可以认为反应停止。——译者注

宇宙的温度降低到十亿度时，核反应过程就突然启动了，此时宇宙大约膨胀了一百秒。那时，又有少许中子发生了衰变，将中子质子的比例减小到了1∶7。绝大部分中子最终的归宿是形成氦-4原子核，除此之外，还有少量的中子形成了氘、氦-3 和锂。最后，宇宙之中的物质组成是 23%为氦-4、77%为氢，还有极少量的氘（万分之几）、氦-3（万分之几）和锂-7（百亿分之几）。这正是今天我们在银河系内外都能观测到的元素丰度。于是，这个核物理学在宇宙最初几分钟中的简单应用，被天文学观测漂亮地证实了（图6.7）。

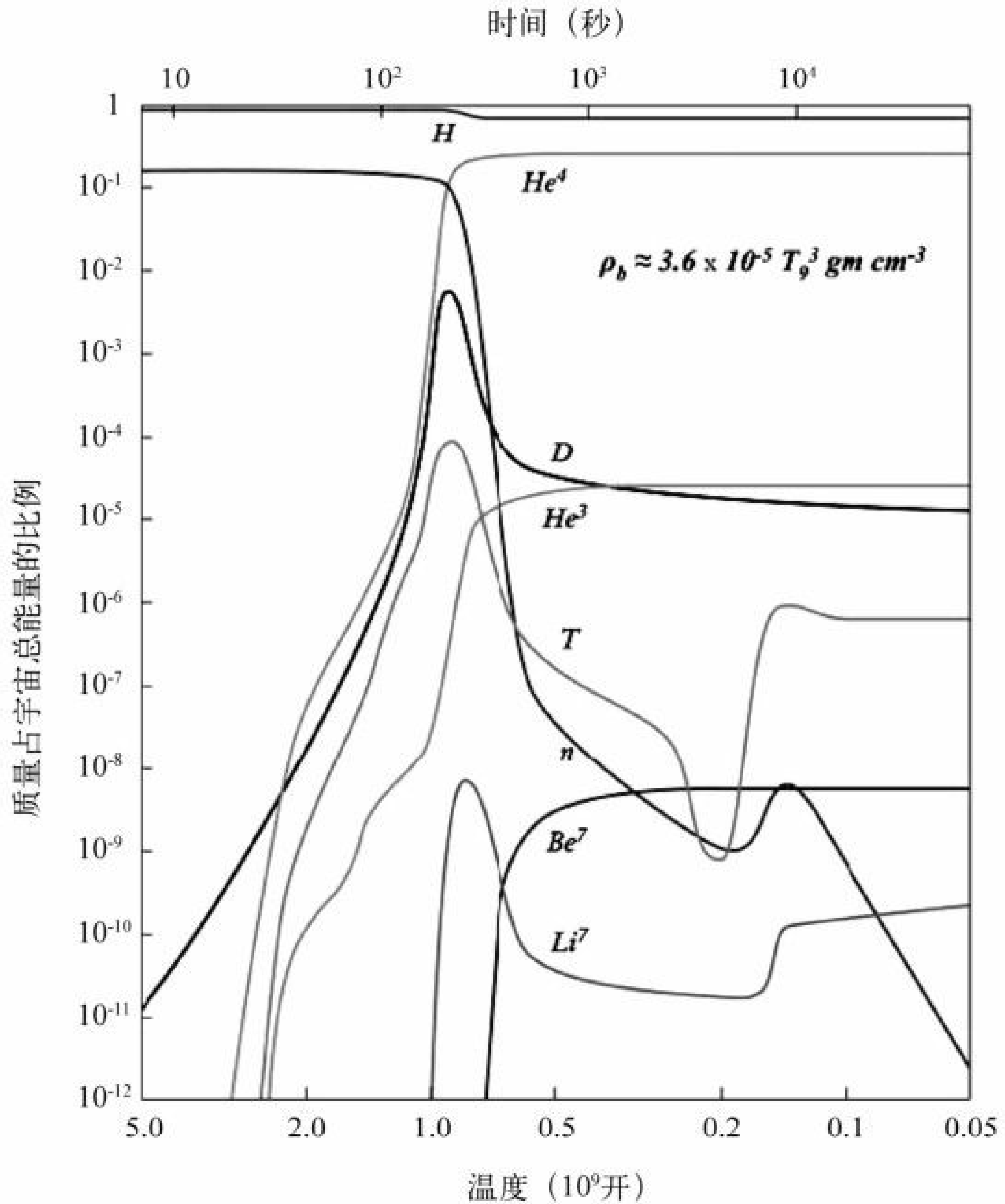


图6.7 随着年龄的增加、温度的下降，宇宙中不同元素所占比例的变化图。不同曲线表示宇宙最初三分钟产生的最轻的化学元素［氢（H）、氦-4（He^4）、氘（D）、氦-3（He^3）、氚（T）、铍-7（Be^7）和锂-7（Li^7）］丰度。标记为n的曲线表示没有被原子核束缚的自由中子会发生衰变，丰度不断降低

从 1975 年开始，随着我们对高能物理和基本粒子的了解越来越深入，重构宇宙遥远过去的状态，成为了一个新的研究方向，叫做“粒子宇宙学”。它注定将彻底颠覆我们对早期宇宙的认识。

注释

[1] F. Hoyle,*The Nature of the Universe*, Blackwell, Oxford (1950) pp. 9–10,基于 1949年他在BBC的演讲。

[2] A. S. Eddington,*The Nature of the Physical World*, Cambridge University Press, Cambridge (1928), p. 85.

[3] R. J. Pumphrey and T. Gold,*Nature* 160, 124 (1947); R. J. Pumphrey and T. Gold,Proc. Roy. Soc. B 135, 462 (1948); T. Gold,Proc. Roy. Soc.B 135, 492 (1948).

[4] 这种叫法是丘宏义在1964年5月的《今日物理》杂志上提出的，他写道：“到目前为止，‘类星射电源’（quasi-stellar radio sources）这个冗长的名字被用来描述这种天体。因为我们对这种天体的性质一无所知，所以无法给它起一个适当而简短的学名，使人们能够望文生义，了解它们的本质。简便起见，在本文中将使用‘类星体’（quasar）一词作为缩写。”

[5] H. Bondi and T. Gold, ‘The Steady-State Theory of the Homogeneous Expanding Universe’,*Mon. Not. Roy. Astron. Soc.* 108, 252 (1948).

[6] 德希特宇宙的尺度因子的膨胀行为是 $a=exp(H_0 t)$，其中H_0是宇宙恒定不变的膨胀速率。

[7] F. Hoyle, ‘A New Model for the Expanding Universe’,*Mon. Not. Roy. Astron. Soc.* 108, 372(1948).

[8] 20世纪 80年代，德希特宇宙成为了暴胀宇宙理论的基础，人们重新发现了德希特空间的稳定性，并称之为宇宙无毛定理。

[9] http://www.aip.org/history/cosmology/ideas/ryle-vs-hoyle.htm.

[10]关于大爆炸理论和稳态宇宙理论的对垒历史，以及天文学观测在其中起的关键作用，更详细的内容参见：H. Kragh,*Cosmology and Controversy: The Historical Development of Two Theories of the Universe*, Princeton University Press, Princeton NJ (1996), chapters. 4–7.

[11]当然，也有一些英国科学家发出了强烈的声音，反对人们沉迷于宇宙学原理。其中最为直言不讳的就是赫伯特·丁格尔（Herbert Dingle，1890～1978）了。他曾经担任皇家天文学会的会长，并曾大力反对爱因斯坦的狭义相对论。他敦促那些“宇宙狂躁症”的患者就事论事，“把锄头当成锄头，而不是什么完美的农学原理”。关于宇宙学的这场哲学层面的争论，更多内容可参见：‘Cosmology: Methodological

Debates in the 1930s and 1940s’, *Stanford Encyclopedia of Philosophy*, http://www.seop.leeds.ac.uk/ entries/ cosmology-30s/.

[12]物质粒子和反物质粒子都携带一定量的重子荷，相应的守恒定律叫做重子数守恒。具有讽刺意味的是，现在人们不再认为自然界的重子数是守恒的了。事实上，如果自然界的各种相互作用都能统一地被一种“万有理论”所描述的话，重子数就不可能守恒，因为这会导致夸克和轻子（例如电子）无法相互转化。

[13] A. Einstein,*The Meaning of Relativity,* Routledge, London (2003), p. 132.

[14]引自写给赫伯特·鲁德的信，转引自：H. J. Rood, ‘The Remarkable Extragalactic Research of Erik Holmberg’,*Publ. Astro. Soc. Pacific* 99, 943 (1987).

[15]哲学家伊曼努尔·康德最先注意到了空间维度和引力衰减程度之间的关系。他指出，牛顿引力之所以服从平方反比定律，是因为空间是三维的。如果空间是n维的，那么随着距离r的增加，引力会按照$1/r^{n-1}$的规律衰减。

[16]今天，我们知道从前的星系之间会经常狭路相逢，发生碰撞。我们所见到的星系之所以形状千变万化，尺寸大小不一，紧密接触和碰撞起到了重要作用。当时霍尔姆伯格还不知道这些，不过他怀疑星系之间可能会发生紧密接触，并留下可观测的痕迹。

[17]霍尔姆伯格已经对问题作了简化，他假设星系是扁平的二维结构，还假设所有相互作用也都发生在同一平面上（摆放模型的桌面上）。

[18] E. Holmberg, *Astrophys. J.* 94, 385 (1941); H. J. Rood, ‘The Remarkable Extragalactic Research of Erik Holmberg’, p. 921.

[19]技术的发展使得人们对天文学数据的分析和理解也变得比以前形象得多。在天文学研究中，随着个人电脑的日新月异，图片和影像也变得极为重要。关于这个问题以及科学研究中影像技术的完整发展历史，参见：J. D. Barrow,*Cosmic Imagery: Key Images in the History of Science*, Bodley Head, London (2008).

[20] K. Waterhouse,The Passing of the Third-Floor Buck,Michael Joseph, London (1974).

[21]邦迪和利特列顿当时不知道爱因斯坦在1924年的提议，即这种电荷的差异也许可以解释太阳和地球为什么会产生磁场。但爱因斯坦很快就放弃了这个想法，因为1925年，皮卡德（A. Piccard）和凯斯勒（E. Kessler）通过实验发现，这种电荷的差异应该小于10^{-20}e，不满足爱因斯坦的要求。

[22] A. M. Hillas and T. E. Cranshaw,*Nature* 184, 892 (1959).

[23] J. G. King,*Phys. Rev. Lett.* 5, 562 (1960).

[24]数学家们也许会对作家、电影制片人 I. A. L.戴蒙德（I. A. L. Diamond）感兴趣。他出生于罗马尼亚，原姓多米尼奇，但后来在美国时改了姓。他经常跟人说自己的名字缩写表示校际代数联赛（Interscholastic Algebra League）。他高中时数学很突出，在 1936～1937年间，曾多次获得美国数学奥林匹克金牌。他就读于哥伦比亚大学，但后来没有继续攻读研究生，因为他撰写的剧本和制作的学生讽刺剧被好莱坞相中了，而且他们提供了一份合约。

[25]在此强烈推荐他的自传：G. Gamow,*My World Line: An Informal Autobiography*, Viking, New York (1970).可惜他 1968年去世时，自传尚未完成，令人非常遗憾。

[26]苏联宇宙学家雅克布·泽尔多维奇（Yakob Zeldovich，1914～1987）曾说，伽莫夫没有回苏联，结果导致其他的苏联物理学家都很排斥他，因为在此后很长一段时期内，苏联官方都竭尽全力阻止任何物理学家出国。参见：Rem Ruffini, ‘Moments with Yakov Borisovich Zeldovich’, http://arxiv.org/abs/0911.4825, p. 2.

[27] G. Gamow,*Physical Review* 74, 505–506 (1948).

[28] R. A. Alpher and R. Herman,*Nature* 162, 774 (1948).

[29] R. H. Dicke,*A Scientific Autobiography* (1975),未发表，收藏于美国国家科学院。

[30] R. A. Alpher and R. Herman,*Physics Today*, 24 August 1988; R. A. Alpher and R. Herman,*Genesis of the Big Bang*, Oxford University Press, Oxford (2001).

[31]这封信见于彭齐亚斯的一篇文章中，参见：F. Reines (ed.),*Cosmology, Fusion, and Other Matters*, Colorado Associated University Press, Boulder, pp. 29–47 (1972).承蒙彭齐亚斯允许，在此处复制。

[32] F. Hoyle and R. J. Tayler,Nature 203,1108 (1964).

第7章
宇宙，如实描绘

我是一只独来独往的猫，对我来说，所有的地方都差不多。

——鲁德亚德·吉卜林（1865～1936，英国作家），《远古传奇》

湍急的宇宙

如果我死了，上了天堂以后，有两件重要的事情希望能搞明白。一件事是量子电动力学的方程组，另一件事是湍流扩散，是关于流体运动的。我对第一件事的预期相当乐观。

——霍勒斯·兰姆（Horace Lamb，1849～1934，英国应用数学家）[1]

爱因斯坦的理论导致早期的宇宙学研究一路突飞猛进，除了极个别的例外，这些研究大都集中在平滑、均匀、膨胀速率保持各向同性的模型上。有两个充分的理由值得人们保持这个倾向。一方面，这个假设非常符合从天文望远镜中所观测到的景象。没有证据能够表明，不同方向上观测到的宇宙膨胀速率有明显的区别，而且天空中的星系分布也和随机的分布差不多。想当年，测量遥远星系的红移大小可是一件极为耗时的工作。在望远镜旁劳作一整夜，也就能测出几个星系的红移。这个情况一直持续到了20世纪70年代。当摄像系统中应用了基于CCD技术的光度检测器以后（此项技术获得了2009年诺贝尔物理学奖），人们就可以迅速完成大量星系的红移测量。现在，有一些大规模勘察星系分布的项目，打算通过测量星系的红移（进而推算出星系的距离），建立一个星系分布的三维立体图。据此建成的三维立体图让人们大吃一惊，原来天上的星系已经形成了线、面之类的结构，进而结成了一张宇宙的“大网”。人们平时看到的星系投影带有一定误导性，根本体现不出星系分布的五彩缤纷和精彩绝伦。另一方面，在这个假设下建立宇宙模型，就能大大简化爱因斯坦方程组，使得求解变得非常容易。[2]

我们前面已经看到，如有必要，可以引入某种“原理”，如宇宙学原理和完全宇宙学原理，以便给这些简化问题的假设提供哲学根据。但也有更充分的理由要求我们更进一步，试着去更加真实地描述宇宙。毕竟宇宙真实的物质分布状况是，其中存在星系之类的结构，也存在其他的不规则性（irregularity）。这些不规则性是从哪儿来的？为什么他们的大小形状是现在这样的呢？尽管各个方向上的膨胀速率差别不大，但不可能是零，因为我们不可能生活在这样一个宇宙中的特殊位置。如果空间都不均匀的话，那么站在空间的每一点上向四周看，也不可能是各向同性的。

勒梅特是第一个认真思考了星系起源的人。栗弗席兹的论文研究了膨胀宇宙在微扰作用下的稳定性，也发现了一个重要的线索：为什么宇宙中像恒星和星系一样的物质“结团”如此众多。如果某个区域中包含的物质略微超过了平均值，它们就会吸引附近更多的物质，把四周变得一

片空旷。这个过程叫做引力不稳定性，是牛顿首先发现的。1902年，詹姆斯·金斯爵士（James Jeans，1877～1946）又对这个问题进行了详细计算。在既不膨胀又不收缩的空间中，这种不稳定性会飞快地增长。但是栗弗席兹发现，如果空间在膨胀，不规则性的增长速率就会变慢。空间的膨胀会试图把相互吸引的物质排斥开，结团的势力和膨胀的势力展开了一场拉锯战。于是，我们很容易理解不规则性的增长速率为什么会变慢。

宇宙学家想知道，这个简单的过程是否能最终解释星系的形成。也许，宇宙“开始”的时候存在着一些微小的“随机”涨落，使空间偏离了绝对均匀的状态。或者也许是后来的一些特殊过程产生了这些涨落，随着宇宙的膨胀，它们越长越大，但最终停止了膨胀，不再变大，而形成了今天我们看到的“星系”结构。这个想法引出了更多需要解释的问题。为什么大多数星系都包含大约一千亿到一万亿颗恒星？为什么它们的形状大都呈特殊的椭球形或螺旋形？而且最重要的是，为什么这么多星系都在旋转？

1944～1951年间，德国物理学家卡尔·冯·魏扎克（Carl von Weizsäcker，1912～2007）开始考虑一种新的可能性。也就是说，宇宙之所以拥有如此丰富的结构，也许是因为它曾经有一段湍流（turbulence）横生的历史。这个想法非常吸引人。湍流是一种常见现象，在生活中普遍存在。打开浴缸的水龙头，你会目睹水流喷涌而出，形成湍流；海浪拍打海岸，或是一辆小轿车绝尘而去时，身后的树叶随风摇曳，你也会在其中发现湍流（图7.1）。然而，会令你吃惊的是，研究流体的数学家认为，这个熟知的生活经验本质上却是无法求解的。如果流体的速度非常快，又产生了很多漩涡的话，你就没办法搞清楚究竟发生了什么。如果流体的速度比较慢，湍流比较温和的话，才有希望搞清楚一些东西。不过，在错综复杂的问题面前，计算机的运算能力还是很快败下阵来。千禧年大奖难题为七个悬而未决的数学题目征集解法，每个答案悬赏一百万美元，其中一个难题正是这个流体力学问题。[3]



图7.1 喷气式飞机引起的湍流

起初，冯·魏扎克打算研究太阳系的起源[4]，以及星系中的恒星运动[5]。不久后，他受到螺旋星系形状的启发，认为这种涡旋状的外观可能是宇宙中曾经出现过的湍流的遗迹，于是他提出，湍流才是理解星系形成的关键所在。[6]魏扎克的老乡沃纳·海森堡也加入了这场科学征程，他是量子力学的奠基人之一，以前也发展过湍流的数学理论。在整个职业生涯之中，他一直对湍流现象非常着迷。[7]

乔治·伽莫夫也被曾经湍流横生的宇宙模型吸引住了。于是，1952年的时候，他发展出了一个相似的理论。[8]不过，星系形成于湍流的想法嘴上说说很容易，要把它转化成一个精确的模型可不容易。实验室中的湍流都还没搞清楚呢，更不用说宇宙中的湍流了，尽管宇宙中的事情至少变化得非常缓慢，而自转物质形成的漩涡也不受边界条件例如浴缸的形状、水龙头和下水口的影响，虽然正是这些边界条件使得浴室里的湍流问题如此复杂。

栗弗席兹已经证明，随着宇宙的年龄逐渐增大，微弱的漩涡都会衰减，变得更加微弱。不过，在热辐射物质主导宇宙的时期，这些漩涡的强度会保持不变。这和花样滑冰的道理一样，如果运动员立在脚尖上旋转的时候把手臂收回来，就会转得更快——如果她把手臂伸出去，就会

变慢。在膨胀的宇宙中，当一个大漩涡的尺寸越来越大时，它的“手臂”也越伸越长，因此，物质的旋转速度也就越来越慢。[9]

湍流有一个特别简单的性质：如果搅动流体时产生了一个大漩涡，这个漩涡的能量就会逐级减小，变成小一号又小一号的漩涡，直到最后由于流体中存在摩擦力，漩涡会全部消失。向一杯水中滴入墨水，然后轻轻地搅动，你就会观察到漩涡逐级减弱的过程。伟大的苏联数学家安德雷·柯尔莫哥洛夫（Nikolai Kolmogorov，1903～1987）在 1941年就曾提出，当漩涡刚出现的时候，它的尺寸最大；在流体黏度的作用下，当漩涡消失的时候，它的尺寸最小；当尺寸处于两者之间的时候，漩涡总是以相同的速率将能量逐级传递下去。[10]因此，在相当宽泛的尺寸范围内，湍流的旋转速度都应该正比于其直径的三次方根[11]，而不论这个湍流是如何产生的，也不论它将来变小后会如何被摩擦力消灭掉。这是一个有趣的结论，因为这样就可以把湍流速度和尺寸的变化关系，跟大小不同的星系旋转速度联系起来。然而，湍流宇宙的演化历史比理论说的复杂得多。

在宇宙最初的三十万年里，声音传播的速度非常快①，漩涡的运动速度比声速低。但是当质子和电子相互结合形成中性原子以后，光子不再与电子发生相互作用，声速也就急剧地减小了。突然之间，涡流的转动都变成了超音速运动，巨变就这样发生了。宇宙中形成了巨大的震荡波，积累起巨大的不规则性，进而影响了物质的分布。星系应该就是在这种震荡波激起的物质漩涡中形成的。我们期待一切尘埃落定之后，就能看见巨大的旋转星系。

① 当时宇宙中的声速大约是光速的$1/\sqrt{3}$。——译者注

这个图景提供了一个解释星系形成的思路，也许旋转的星系源自于特定的柯尔莫哥洛夫型湍流。不过，我们得先确定宇宙湍流的起源，同时也要搞清楚超音速物质的积累规律。

当时没有多少人支持伽莫夫的湍流宇宙研究，于是这个理论就被搁置了下来。直到 20世纪 60年代中期，广岛的成合秀一（Hidekazu Nariai）和莫斯科的里奥尼德·奥泽诺（Leonid Ozernoy）分别领导的两个厉害的研究小组才开始发展这个理论。在1964～1978年间，湍流宇宙一直被当作是星系起源的候选理论，人们向其中倾注了大量精力。从宇宙学研究的角度看，这个理论令人难以捉摸，因为要想从爱因斯坦方程组中找出一个能够描述真实的湍流宇宙的解，实在是太难了。你所能做的事情，只有是寻找宇宙早期形成的湍流留下的蛛丝马迹，看它们是否导致物质的密度和压强出现不规则性，使之聚集为旋转的物质孤岛，并

最终变成现在我们从天空中看到的星系。

但后来遇到的两个重大问题将这个思路完全堵死了。起初人们以为，这些漩涡不会对宇宙膨胀的过程产生太大的影响。唉，可惜这种说法被证明是错的。当你逆着时间回溯时，尽管漩涡的速度保持不变，可它们产生的引力效应对膨胀宇宙的影响却是越来越大。最终人们发现，必须存在这样一个时期，整个宇宙混沌未分一团糟。为了能够形成今天我们所看到的星系，这些漩涡必须旋转得足够快，这就要求宇宙的混沌时期离现在非常近，但这会和微波背景辐射的观测结果相矛盾。这个理论同时也会破坏宇宙诞生大约三分钟时发生的氦元素的合成过程，因而与我们观测到的化学元素丰度相矛盾。[12]

湍流机制的第二个重大问题是，我们对旋转星系的认识发生了根本的改变。直到1974年，天文学家们还认为所有的椭圆星系都在旋转，而且转得越快，形状就越扁。但后来，牛津大学的研究生詹姆斯·宾尼证明，解释椭圆星系的形状为什么是扁平的，并不要求星系必须在旋转。[13]这些星系中的大多数恒星都沿着随机的轨道运动，而星系的总体形状是在它产生的过程中形成的。渐渐地，这种想法得到了天文观测支持：许多椭圆星系的旋转速度都太慢了，不足以导致这种扁平的形状，而一些椭圆星系甚至是绕着“错误的”对称轴旋转。这两个进展导致了星系形成自湍流的理论的迅速消亡。事实上，湍流对宇宙膨胀产生的影响太过剧烈，会导致极其严重的不规则性和扭曲，而这恰恰与我们今天观测到的各向同性、均匀的宇宙大相径庭。

畸变的宇宙：从Ⅰ号到Ⅸ号

任何宇宙演化理论的目的，都是试图寻找尽可能简洁的宇宙诞生时的初始条件。当我们将已知的相互作用力置于这些初始条件之上时，世界上一切复杂现象就可以统统冒出来。

——乔治·勒梅特[14]

直到1950年，宇宙学研究的主要内容还是那些到处都一样、各个方向的膨胀速率也都相同的宇宙学模型。这叫做“均匀”（到处都一样）和“各向同性”（各个方向都相同）的宇宙。它们是对称性最高的宇宙学模型，也是爱因斯坦方程组最简单也最容易理解的解。时不时地，也会出现另一种类型的宇宙模型，例如由卡斯纳、爱因斯坦和罗森、施特劳斯、托尔曼和哥德尔发现的宇宙模型，或者最先由乔治·勒梅特和伊夫金·栗弗席兹提出的、在各向同性均匀的基础上有一些微小扰动的宇宙模型。然而，这些不同寻常的宇宙模型能被发现纯属偶然。人们并不清楚这样的模型一共有多少种，应该如何进行分门别类。这种状况在1951年的时候发生了改变，当时亚伯拉罕·陶伯（Abraham Taub，1911～1999）完成了一项意义深远的系统性研究，囊括了所有均匀但不是各向同性的宇宙模型。[15]无论身处这个宇宙的哪个角落，你都会看到相同的宇宙演化历史。天文学家就像一只独来独往的猫，无论他们身在何处，

宇宙的任何角落看起来都差不多。但如果朝着不同的方向看，你就会发现一些不同之处。爱德华·卡斯纳在20世纪20年代发现了第一个这样的宇宙模型，那是其中最简单的一个。可是，这种均匀但不是各向同性的宇宙一共会有多少种呢？

1977年，我在加州大学伯克利分校做博士后研究，当时亚伯拉罕·陶伯（大家都叫他“亚伯”）是数学系广义相对论研究组的组长。我属于天文系，但我会参加陶伯小组的组会和非正式的午间聚餐。对年轻的科学家来说，陶伯相当令人生畏，因为他的知识极为广博，经验极其丰富，而且非常挑剔，从不与人聊家常：你可不能跟他讨论不成熟的想法。他建立了伯克利的计算机实验室，同普林斯顿的冯·诺依曼有合作，也同爱因斯坦和哥德尔有联络；他是震荡波和流体力学的世界级权威；在罗伯特逊的指导下，他于1935年取得了博士学位，因此也是广义相对论和宇宙学领域的专家。他管理的组会秉承了传统的法官审问式风格（inquisitorial style），如果他还有疑问没解决，就从不允许主讲者继续往下讲；在学术讨论中，就连来访的诺贝尔奖得主也被他当作研究生一样对待。起初，他看起来脾气很暴躁，但是渐渐地你会发现，他还挺慈爱的：他对待年轻的学者，就像是一个父亲，管理着一个巨大的家庭，孩子们随时会任性胡闹，要费尽心思使他们步入正轨。

1951年，陶伯发现，早在 1898年时，意大利伟大的几何学家路易吉·比安基（Luigi Bianchi，1856～1928）就已经搞清了所有可能存在的、对任何地方的观测者来说都一样的空间。[16]如果不同方向之间存在差别，这些差别也不应该随着地点的变化而变化。比安基的一生都在比萨市的几个优秀的数学研究所中担任教授。他证明，这样的空间只有九种，分别用罗马数字Ⅰ～Ⅸ表述。



图7.2 路易吉·比安基

按照比安基的分类方法，陶伯找到了与膨胀宇宙相符的所有类型的均匀空间。[17]罗马式阶梯上陈列着比安基归纳出的Ⅰ号到Ⅸ号的空间类型，从最简单的平坦的欧几里得几何，如卡斯纳的宇宙，到黎曼和罗巴切夫斯基的各向同性的弯曲空间，以及几种新型的并非各向同性的弯曲空间。随着宇宙的膨胀，其中一些空间甚至可以改变曲率的大小：在大部分时间里，它们就像负曲率的“开放”宇宙，但突然之间，又会变成正曲率的空间。又有一些空间，就像哥德尔的宇宙那样，能够旋转，但还有一些空间能以一种扭曲的方式一边旋转一边膨胀。

陶伯没能将所有这些可能性都从爱因斯坦方程组中解出来，但是他将其中一部分解了出来，并且把卡斯纳的宇宙以一种我们前面见过的简洁形式写了下来。直到今天，我们仍然无法从方程组中解出其中最复杂的那个宇宙，但在过去的50年中，几乎所有类型宇宙的总体行为都逐渐得到了定性描述。20世纪60年代，人们对这些“比安基的宇宙”进行了大量研究。直到80年代末期，这种模型才退出了人们关注的焦点。通过研究它，人们可以理解为什么天空中的微波背景辐射是如此地符合各向同性。

比起弗里德曼、勒梅特和德希特的最简单的宇宙，陶伯的各向异性的膨胀宇宙模型要复杂得多，也诡异得多，因为它拥有四个关键性质：

1.剪形畸变

2.旋转

3.存在相对于径向哈勃膨胀的运动

4.曲率各向异性

第一个性质的意思是，拿起一个圆形的球，用手捏它的上下两侧，将它变成鸡蛋的形状。如果说各向同性的膨胀宇宙是由一系列体积增大的圆球所组成的，那么剪形畸变的宇宙就是由一系列体积增大的椭球所组成的（图7.3）。宇宙既要膨胀又要旋转时，也必然会出现剪形畸变（图7.4）。几种性质的综合作用也很容易理解。第三个性质更加微妙。想象一个膨胀的圆球，有一种观察膨胀效应的方法是，从球心沿着半径向外画一条直线。通过测量这些线段的长度，我们就可以推断出宇宙膨胀了多久（图 7.5）。我们倾向于假设我们正在沿着这些虚拟的线段运动（其他星系中的每一个观测者都这么想）。但是我们不能断定，所有天体都随着宇宙的整体膨胀而运动——宇宙学家们将宇宙的总体膨胀称为“哈勃流”。我们所在星系团也许正朝着另外一个方向运动，或者其中存在着大量的随机运动。如果这些背离哈勃膨胀的运动相对缓慢，就像地球围绕太阳转，又和整个太阳系一起围绕银河系中心转那样，或

者参与这类运动的物质不太多的话，就不会对宇宙的总体膨胀产生任何可观测效应。但如果这类运动牵扯到大量物质，相比哈勃流的速度又很高的话，就会对宇宙的膨胀产生影响，引起一种剪形畸变。

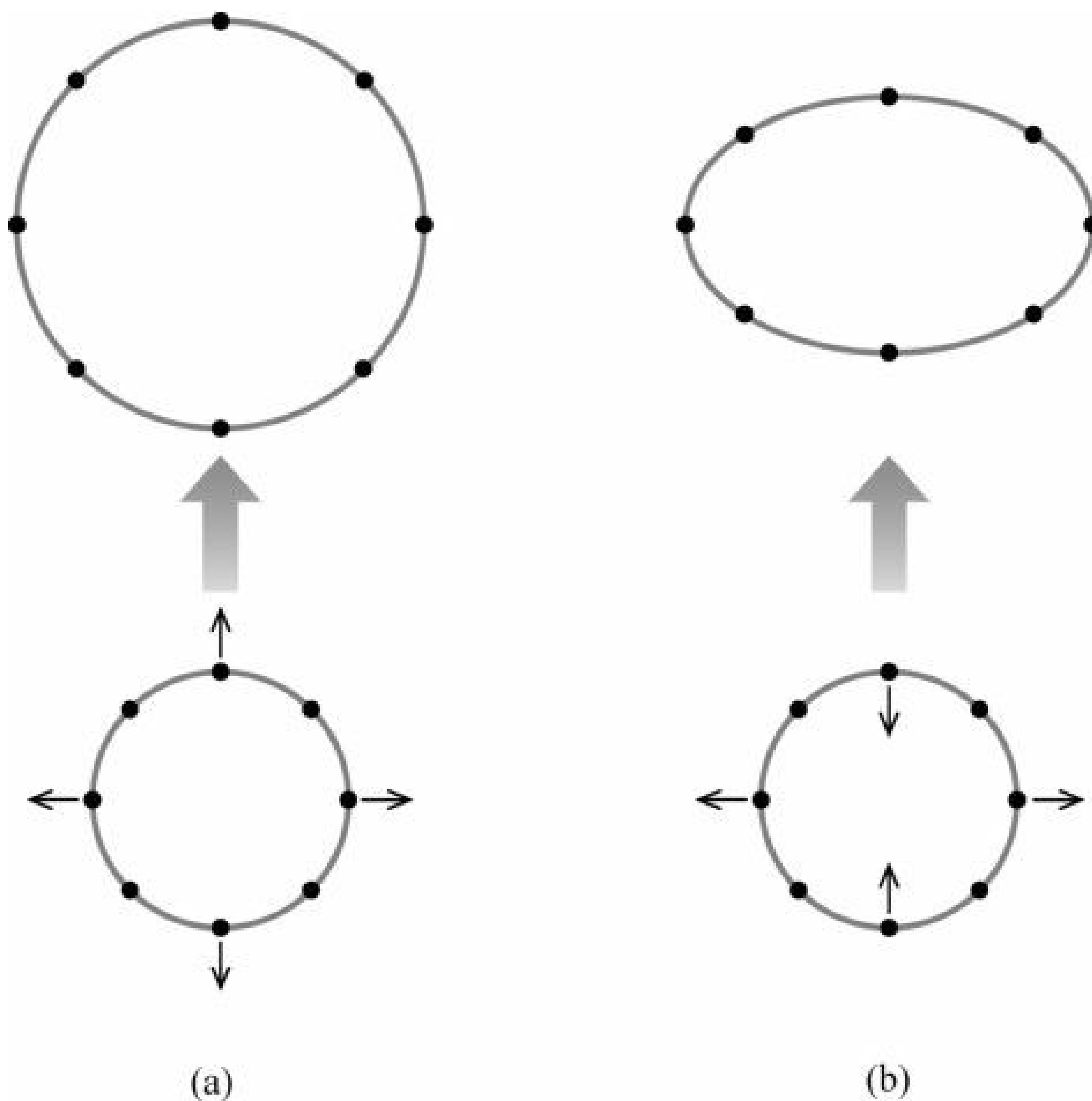


图7.3 剪应力把球体扭曲成椭球体：(a)没有剪应力时的膨胀，(b)发生了剪形畸变的膨胀

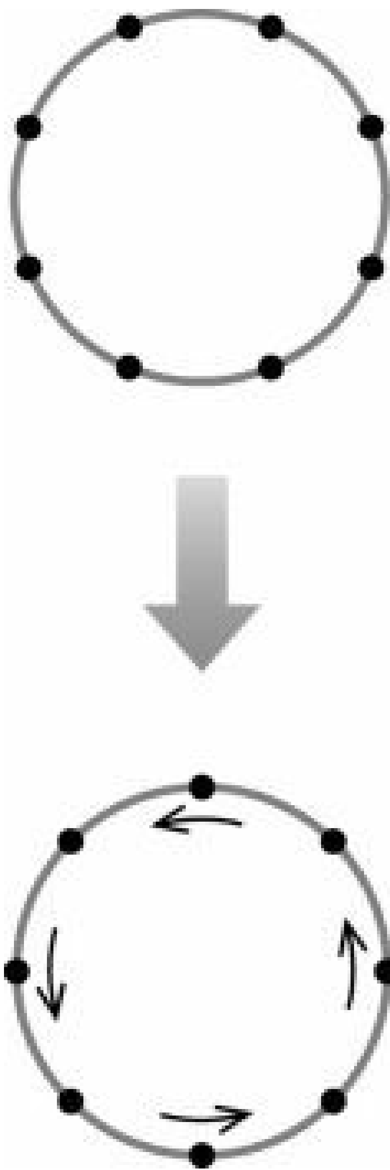


图7.4 旋转对球体的扭曲作用

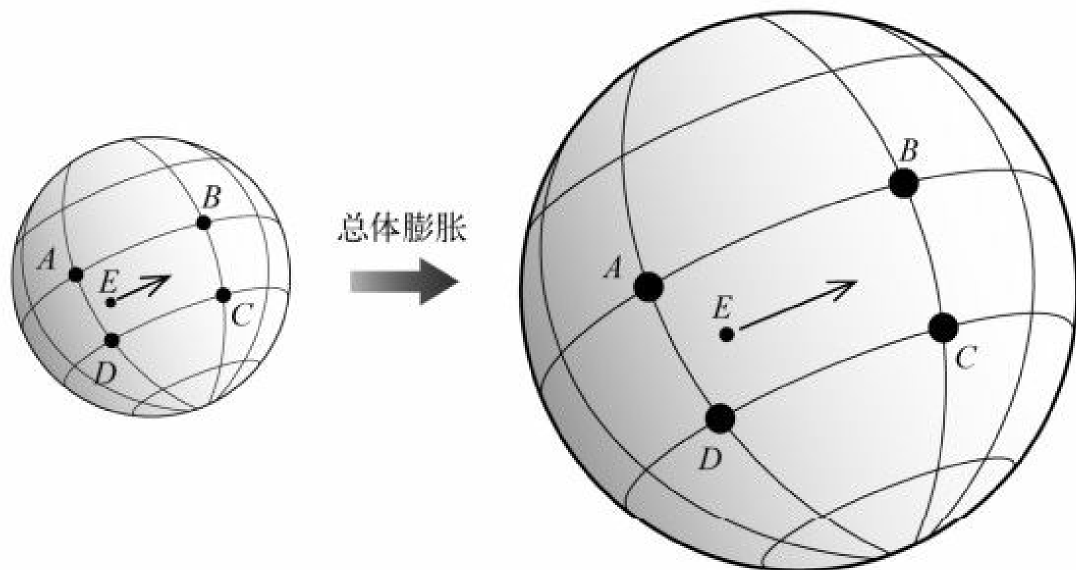


图7.5 在一个各向同性的膨胀宇宙中，有一个星系相对于总体膨胀还有一个额外的运动

打一个简单的比方就清楚了。假设我们的膨胀宇宙像一个气球的表面，只有两维空间。如果把睡觉的蚂蚁放在气球上，然后把气球吹大，这些蚂蚁就以均等的速度相互远离。蚂蚁们会随着它们的宇宙——气球的表面而膨胀。但是如果蚂蚁清醒过来，随着气球的膨胀开始到处乱爬的话，它们之间就额外产生了一种相对的运动，这种运动不能简单地解释为宇宙的整体膨胀。如果太多的蚂蚁朝着一个地方爬，就会在气球的橡胶表皮上弄出一处凹陷，于是就改变了这个方向的膨胀速率。

我们知道，爱因斯坦理论的精髓在于，物质和能量的运动会改变空间的几何性质，其中包括曲率的大小。在弗里德曼和勒梅特发现的各向同性的简单宇宙中，空间的曲率和膨胀速率一样，都是各向同性的——从每个方向看都相同。还有一类陶伯没有算出来的宇宙，它们的性质完全不同，因为其中的空间曲率也会随着方向的不同而变化。卡斯纳的宇宙是最简单的一种膨胀速率为各向异性的宇宙，但它的空间几何任何时候都是平坦的，所以曲率是各向同性的。实际上这只是一种特殊情况。比安基空间最复杂，在任何时候的膨胀速率和空间曲率都是各向异性的（图7.6）。

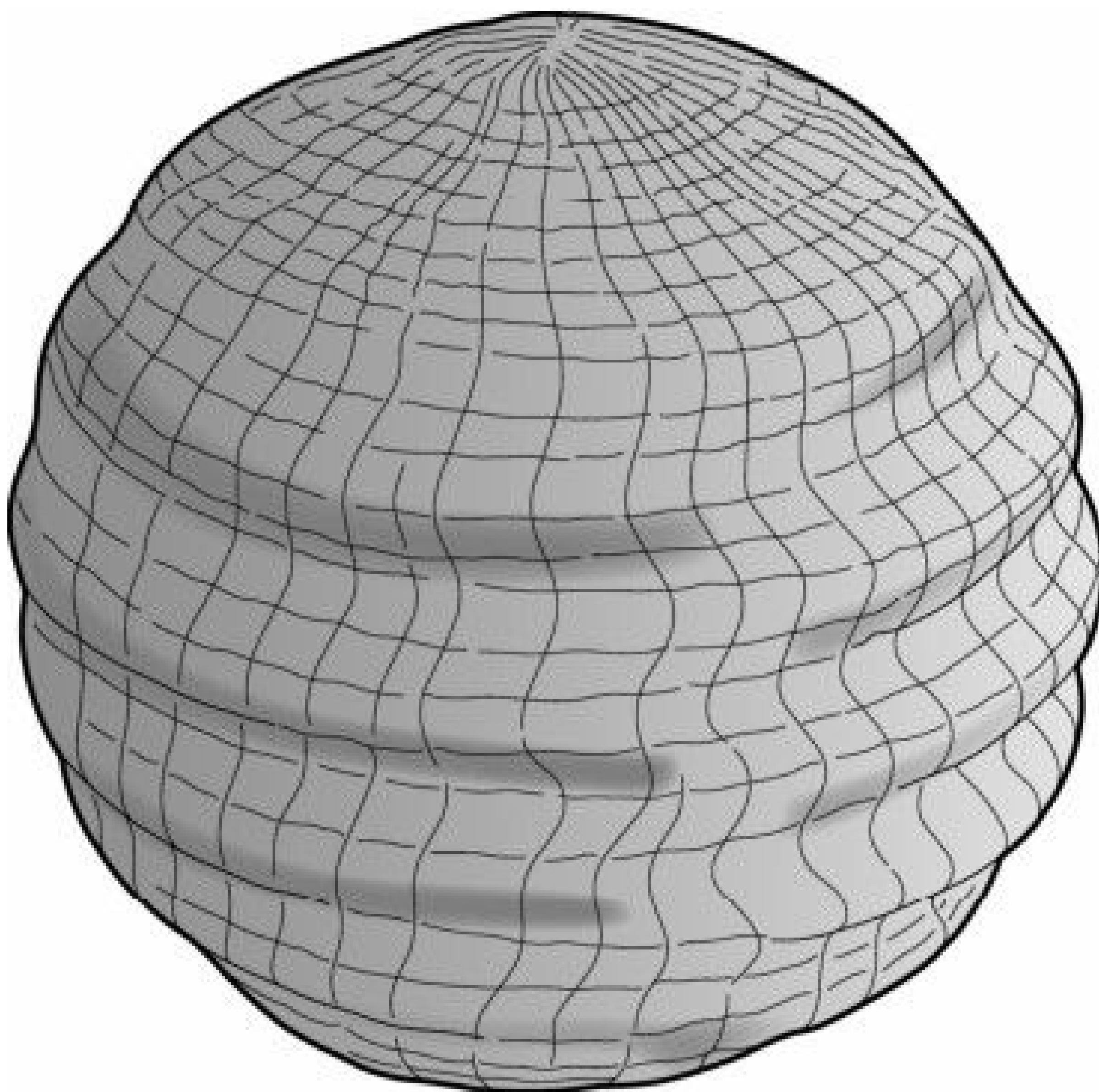


图7.6 曲率各向异性的空间，有的方向上看像是一个开放的宇宙，有的方向上看像是一个闭合的宇宙。随着时间的流逝，这些方向上的曲率可以从正变到负，反之亦然

这些新型宇宙的曲率是各向异性的，看起来就像往各向同性的宇宙中加入了引力波。不同强度的引力波朝着不同的方向传播，于是沿着传播的方向产生了不同的曲率。如果又想让这些复杂的事物能够同时存在，又想让宇宙看起来是均匀的，那么就得给宇宙加上了一组很强的限定条件，这就是为什么这种类型的宇宙数量非常少。这个问题最先是由比安基发现的。如果允许宇宙的不同角落看起来并不相同的话，那么这样的情况就有无数多种。陶伯对宇宙进行分门别类，就像艺术家为自己

的创造力套上一个画框，又像雕塑家将自己的思想集中表达在一块石头上一样。自发引入的限制能令人全神贯注。

均匀的宇宙和一扇新的观察窗口

世界总是黑暗的，光明只不过是把黑暗隐藏了起来。

——丹尼尔·K.麦基南

我们在上一章说到，一个偶然的发现震惊了世界，那就是早期的炙热宇宙留下的微波背景辐射。微波背景辐射的平均温度大约是 2.7 开，但是普林斯顿的射电天文学家很快发现了一种方法，能够更精确地测定这个温度的大小。想要测出温度的精确数值很困难，因为我们首先需要找到一个绝对准确的参考温度，以便测量时比较两者的大小。今天，测量温度的最先进的方法是将一个人造卫星发射到太空中，但是人造卫星必须携带一小瓶温度约为 2.7 开的液氦以作为参考标准。最后，液氦会全部蒸发掉，这也就限制了实验所能持续的时间。

1967年，戴维·威尔金森（David Wilkinson，1935～2002）和布鲁斯·帕特里奇（Bruce Partridge）应用了一种由罗伯特·迪克发明的精巧电子装置，能够以很高的精度测量天空中两个不同方向之间的温度差别，比测量平均温度的精度高得多。差分测量能够得到很高的精度，是因为不需要测量温度的绝对大小。威尔金森和帕特里奇测量了微波背景辐射的强度变化，在探测器的误差范围内，他们什么结果都没发现。于是，他们证明了天空中不同方向的背景辐射的温度之间的差别小于0.1%：微波背景辐射体现出无与伦比的各向同性，宇宙中的物质结团没有强大到足以让宇宙膨胀产生明显畸变的程度。[18]

这个发现令世人皆为惊叹。这说明宇宙膨胀的过程极为符合各向同性；实际上，太符合各向同性了，宇宙学家开始考虑转变态度，重新审视在宇宙中看到的高度对称性。在此之前，宇宙学家们已经假设了宇宙非常接近各向同性均匀的状态。最大的问题在于，如何才能解释为什么光滑的背景之上存在一些微小的不规则性，以及这些微小的不规则性是如何变大，变成我们今天看到的星系和星系群的。微波背景辐射的高度各向同性发现后不久，宇宙学家又开始将背景的均匀性和宇宙膨胀几近完美的各向同性当作重大的谜团。毕竟，如果我们随机地从爱因斯坦的方程组中选择一个解，或者从比安基的空间陈列馆中选择一种空间，选出不太规则的、各向异性宇宙的概率要比选出美丽而均匀的、各向同性宇宙的概率大得多。既然有那么多方法能让宇宙变得不规则，为什么我们的宇宙却如此小众地体现出均匀和对称呢？

混沌的宇宙

条条大路通罗马

——让·德·拉封丹（1621～1695，法国诗人）[19]

1967 年，受到微波背景辐射观测结果的激励，马里兰大学的查尔

斯·米斯纳（Charles Misner）提出了一种全新的想法。[20]不再假设宇宙是均匀而各向同性的，这样是从有序中产生有序，我们为什么不去证明有序也可以从混沌中诞生呢？

米斯纳的“混沌宇宙”计划试图证明，爱因斯坦方程组存在这样的性质，无论宇宙开始膨胀之时有多么混乱，如果等上足够长的时间（而且到目前为止，我们有将近一百四十亿年的时间可以用来等），宇宙就会变得越来越均匀，越来越符合各向同性。

这个想法非常吸引人，而且有着重要的哲学意义。如果事实真的如此，就意味着我们不需要了解宇宙是怎样诞生的（假如它有个开端的话），就能解释它现在的模样。米斯纳想要证明，如果从一个混沌的状态开始膨胀，宇宙的早期总会产生足够多的摩擦力，能够将不规则性都抹平，最终使宇宙的膨胀变得各向同性而均匀。这就像用力搅动一桶油。你闭上眼睛，看不见油中搅出的漩涡，但这样你也肯定知道，一分钟以后，油的表面又会平静如初。油中的运动最终消失在阻尼之中，油面又恢复了光滑和平静。宇宙也会像油一样吗？

这个混沌宇宙的新理论显得雄心勃勃。在此之前，天文学家都是在爱因斯坦方程组的众多解中寻找一个与我们所见到的宇宙尽量符合的，他们已经对当时的方法非常满意了。那个最简单、最具对称性的解与现实符合得就很好。后来，宇宙微波背景辐射的观测结果越发凸显出一个事实，那个解与现实符合得不仅仅是很好，而且是非常之好。为什么会这样呢？这就是米斯纳提出的问题。一些宇宙学家满意于这样的解释：现在的宇宙拥有高度的对称性是因为它一开始时就是这个样子——“事情之所以现在如此，是因为它们过去就是那样”，托马斯·戈尔德曾如此讽刺道。[21]这根本解释不了什么问题。于是米斯纳想要找出一个更合理的解释。如果他能证明，现在的对称性可以从任何可能的初始条件中产生（或者从其中的大部分可能中产生），那么他对宇宙结构的解释就非常有说服力。

起初，米斯纳的想法看起来很有前途。基于比安基、卡斯纳和陶伯发现的各向异性空间，所能构成的最简单的宇宙确实会随着膨胀的进行，迅速趋向各向同性，而且对称性变得越来越高。不过很可惜，在更复杂一点的情况中，也就是在更可能出现的情况中，并不能找到这样的性质。你不能指望宇宙通过不断地膨胀就能一点一点地抹平其中的不规则性。在一些宇宙模型中，这些不规则性根本不会消失，除非有一些特定的物理机制能够比宇宙膨胀更快地减弱不规则性的强度。米斯纳面临的挑战在于，如何才能找出这样的平滑机制。

不幸的是，这次搜寻一无所获。德州大学奥斯汀分校的理查德·马茨纳和我证明，微波背景辐射本身的性质就说明，几乎没有多少不规则

性曾经被抹平掉。[22]任何不规则性的抹平过程都会产生热（多亏热力学第二定律，才使得不规则性携带的能量有了如此归宿），微波背景辐射正是这种热在今天的体现。抹平过程发生得越早，产生的热也就越多。当今宇宙中的总热量（大约是每个原子对应十亿个光子）就大大限制了允许被减弱和抹平的不规则性的总量。事实上，只有非常少量的不规则性在宇宙的早期历史中消失了：如果宇宙起源于一个非常不规则的状态，就不可能利用自身的膨胀将其中的不规则性都平滑掉，只剩下我们现在看到的各向同性空间，同时又不产生大量的热辐射。

米斯纳同时也强调，无论当时人们研究过的宇宙模型是否拥有高度的对称性，其中都有一个不同寻常的特点。光波以有限的速度传播，于是就为任何信号的传播或是抹平机制的传递施加了一个宇宙速度的限制。如果宇宙诞生后过了t秒钟，那么光在这段时间内只能传播一段“视界”的距离约为$10^5 \times t$千米，包含了$10^5 \times t$个太阳的质量，所以宇宙诞生后的10秒钟内，光只能传播100万千米（地球到太阳的距离是1.5亿千米），所影响到的质量大约是太阳质量的100万倍。但在宇宙中，我们所能观察到的最大范围的均匀性是这个的10^{15}倍。

这个结论给混沌宇宙理论投下了一个巨大的阴影。在宇宙甚早期，任何能削弱不同区域的温度、物质密度和膨胀速率差异的机制都被限制在了很小的尺度上，这就无法解释为什么从大尺度上看宇宙的不同区域表现得如此均匀，天空中的微波背景辐射的温度如此一致。

如果你看天上两块相隔超过两度的不同区域（其温度在极高的精度下相等），但在宇宙最初的二十五万年中，光信号都没有足够的时间在这两个区域之间传播，以便抹平其中的能量和温度的差别。此后，光子开始能够在宇宙中自由传播，最终到达了我们的眼睛，我们观测到的微波背景辐射就是当时宇宙的快照。

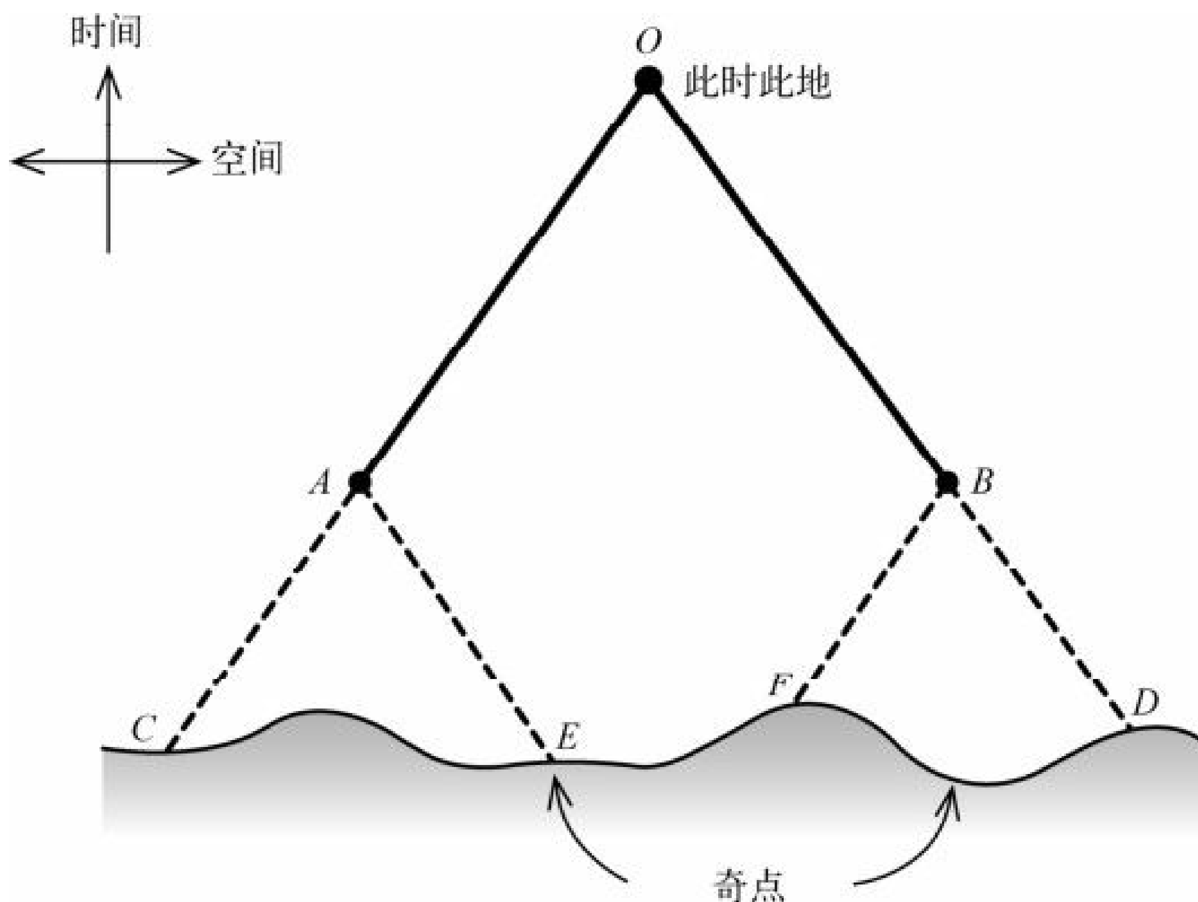


图7.7 如图所示，朝我们传播的光线在时空中形成了一个“光锥”（OCD）①。光锥定义了视界，处于光锥之外的那部分宇宙我们是看不到的。选择过去光锥上的两个时空点（A点和B点），当时微波背景辐射的光子刚刚获得了自由，开始飞向我们的眼睛，而宇宙的年龄只有几十万年。我们可以找到这些时空点所拥有的过去光锥（ACE和BFD）。就算你回溯到宇宙的开端，这两个光锥也不会相交。这说明在对于当时的宇宙来说，光信号还来不及在A点和B点之间传播，因而A点和B点的状态不可能通过任何以光速传播的物理机制而趋于一致。但是人们发现，A点和B点的温度和密度几乎是一样的，其中的差异不超过十万分之一

① 试着把纸上的二维平面想象成时空。水平方向代表空间，垂直方向代表时间，向上代表未来，向下代表过去。光是宇宙中最快的信使，因此，人们在纸上画出光的运动轨迹，用来研究时空的因果性质。任取纸上的一点O，代表某个时刻站在地球上的观测者。穿过O点对称地画两条交叉的直线，代表观测者身上接收和发射的光线，所形成的对称锥形图案就叫做光锥。光锥把观测者周围的时空分成了两部分，光锥

外的一切与此时此地的观测者都没有因果联系。比如，此时此刻，1.5亿公里外的太阳就与我们没有因果联系——我们看不到“此时此刻”的太阳，只能看到8分钟前的太阳。所以与此时此刻的我们有因果联系的是8分钟前的太阳。同理，此时此刻，我们看到的所有事物大都是事物一段时间之前的状态。这些远近不同的物体在不同时刻发射或反射的光同时到达了同一个地方——我们的眼睛，我们就说那时那里的它们位于我们的过去光锥之中，也就是图中光锥朝下的那个开口。——译者注

这个简单推理意味着，要么你得想出某种“新物理机制”，能够让抹平过程不再受到宇宙速度施加的限制；要么宇宙的初始条件比较特殊，但这就排除了那些奇异的不规则宇宙，它们不可能是我们的宇宙。

“搅拌大师”宇宙

谐波混频是世界上顶级DJ使用的一种高级技术……显然，学会如何异调混频（mix in key）是一种高级的DJ技术。如果你开始学习异调混频时水平还不够高的话，没关系，只要你铭记在心，异调混频是你未来成为世界级DJ大师之前所要研习的必修课程。

——《DJ大师课程》[23]

在宇宙的早期，不规则性的抹平会受到视界大小的限制。1969年，米斯纳发现了爱因斯坦方程组的一类新的解，对这个问题作出了回应。[24]这种宇宙的体积是有限的，是比安基发现的空间中形式最复杂的一种。这是一个“闭合”宇宙，从一次大爆炸后开始膨胀，体积达到最大以后就开始收缩，再经历一次大塌缩。然而，这种宇宙的膨胀方式太复杂了，以至于根本不可能求出爱因斯坦方程组的解。不过，我们仍然可以概述其大致行为，并用计算机研究其中的一部分。后来，人们发现这种宇宙的膨胀行为是混沌的，根本无法预测[25]，于是这就成为20世纪70年代末科学界关注的主要研究方向。

根据一种美国商用食品搅拌器的特点，米斯纳将这种类型的宇宙叫做“搅拌大师”（Mixmaster）。[26]他相信，这种宇宙有着高度扭曲、瞬息万变的几何结构，使得光能够环绕宇宙而行，从而将整个宇宙搅匀。可惜进一步的研究表明，尽管光可以在“搅拌大师”宇宙中跑出很远的距离，但这种事情极为罕见，因此整体上搅拌进行得并不充分。

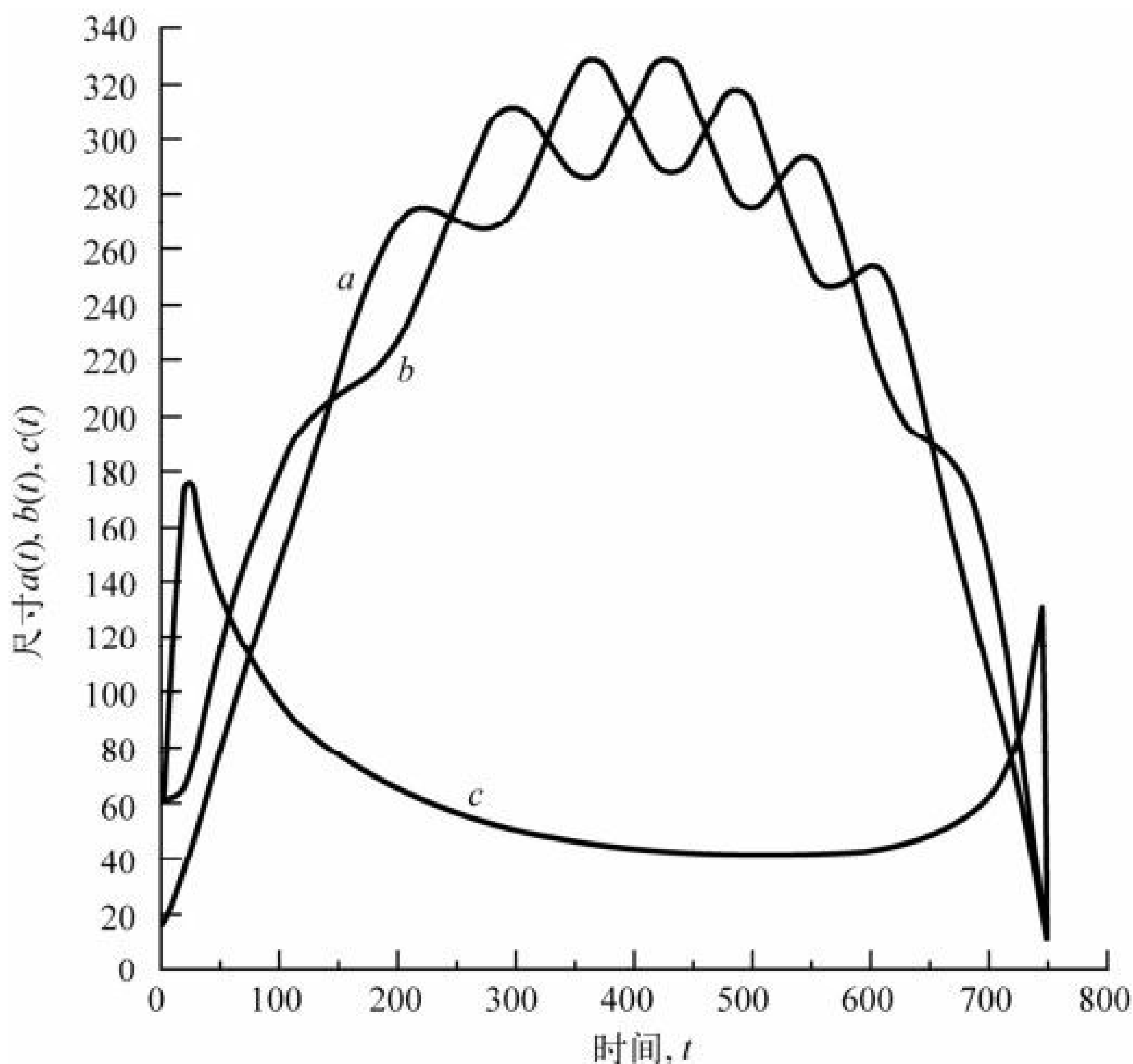


图7.8 “搅拌大师”宇宙的振荡。在宇宙的前半生中，它的体积不断增大，其中有两个维度在膨胀，而第三个维度在收缩。在宇宙的后半生中，整个体积不断缩小，直到变成零，为下一次大爆炸做好准备。哪个维度膨胀、哪个维度收缩完全是随机的，而且总是变来变去，就像一个颤动的果冻。在某个周期里，某两个维度不断发生振荡，而第三个维度稳定地变化；在另一个周期里，原先振荡的维度和原先稳定变化的维度互换了角色。宇宙经历了无穷多次振荡，本图只显示了其中一部分。相互垂直的三个维度的尺寸分别为a(t)、b(t)和c(t)

“搅拌大师”宇宙是人们见过的满足爱因斯坦方程组的宇宙中最复杂的一个（图 7.8）。如果追溯它开始的样子，你会发现它经历过无数次振荡。在其中的任何一个时刻，宇宙都有两个维度在膨胀，一个维度在

收缩，就像卡斯纳的宇宙。然而，哪个维度收缩、哪个维度膨胀却是随机的，换了一轮又一轮。如果追溯到宇宙的开端，你就会发现膨胀和收缩的维度已经转换了无数次。这就像一个颤动的果冻。[27]

这种复杂性完全是由爱因斯坦的引力理论造成的，在旧的牛顿引力理论中根本找不到对应的宇宙模型。你可以设想一个宇宙开始膨胀，不同维度的膨胀速率不同，向其中各个维度都加入引力波，沿着各个维度在空间中掀起涟漪。同时，这些引力波也弯曲了它们周围的空间。向内传播的引力波造成了巨大的空间曲率，最终使得引力波调转了方向；同时，收缩维度也转而开始膨胀，而另外两个维度中的其中之一由膨胀转为收缩。当追溯过去，直到宇宙诞生的那一刻时，你就会发现这样的转换已经重复了无穷多次。

这种奇怪的事情让人想起一个非常古老的哲学问题。公元前5世纪，埃利亚的芝诺提出了一个问题：你能在有限的时间中做无穷多件事吗？芝诺精心构思了一个关于无穷的悖论，想用它来向其他哲学家发出挑战。古时候没有人能回答这些悖论提出的问题。比如说，如果你和门之间的距离是 1米，那么为了走到门口，你必须先走到 1/2米的地方，然后是1/4米的地方，再然后是1/8米、1/16米，依次类推，以至无穷。这些长度一份一份地加起来[28]等于1米。但是芝诺反过来说，你必须走无穷多次才能走到门口，也就是说，你永远也无法走到门口。

当我们的时钟走过一段有限的时间时，“搅拌大师”宇宙已经经历了无穷多次、物理上可区分的振荡。米斯纳解释说，走到门口虽然需要走过无穷多个区间，但这并不是物理上可区分的事件。如果以此为由认为芝诺的悖论在此并不适用，那么你就必须认定“搅拌大师”宇宙的年龄是无限大的，因为当你还没追溯到 $t=0$ 的时刻时，就已经有无穷多次、物理上可区分的振荡发生了。[29]

从物理学的角度讲，这个事情看起来很奇怪，不过数学家已很习惯于此。例如，画$y=x^2\sin(1/x)$的函数图，如果你有一根无穷细的铅笔，你就可以在 x 轴上靠近 $x=0$ 处的任意小区间内画上无穷多个振荡（图7.9）。

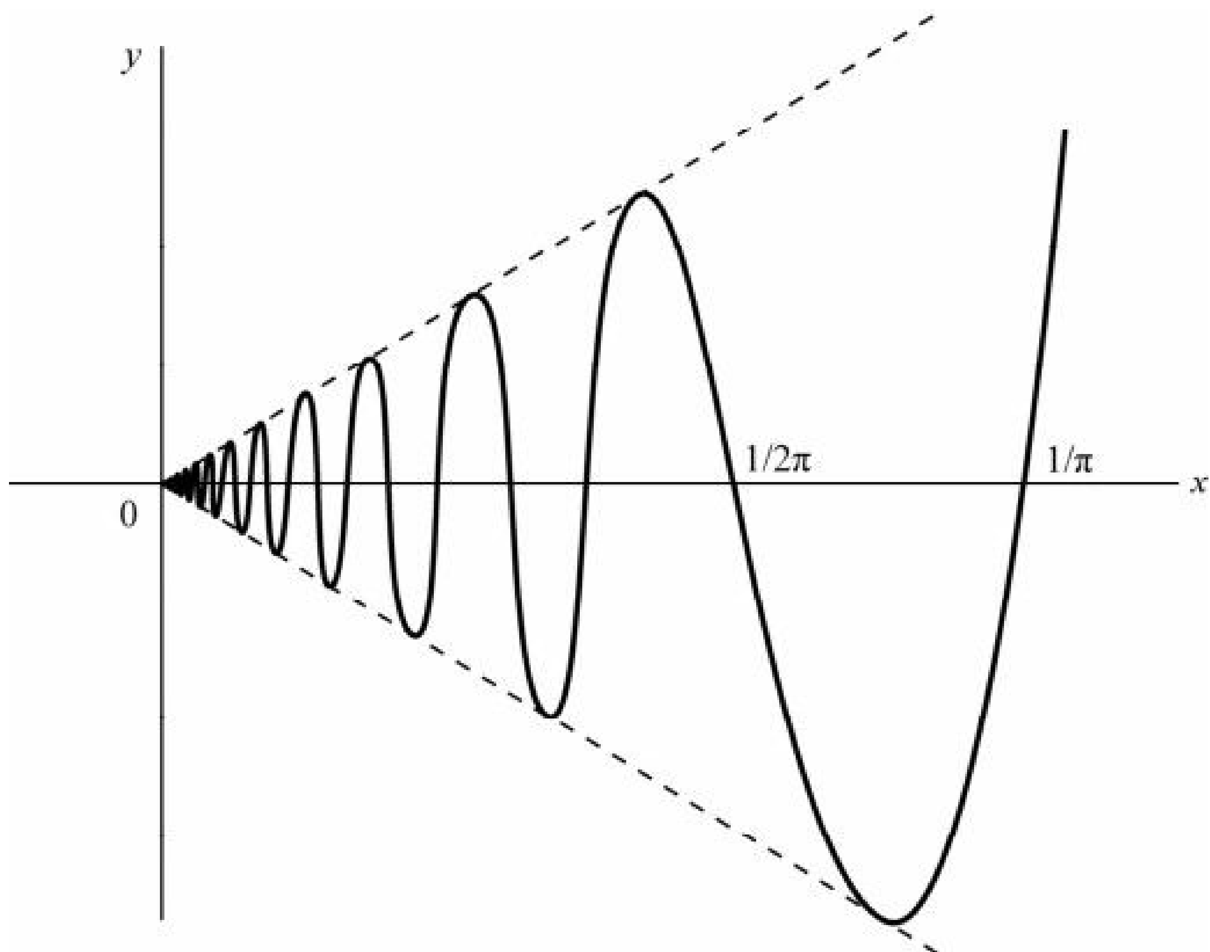


图7.9 y=^{2}xsin(1/x)的函数图。实际上，在它接近x=0之前，会经历无数次振荡。我们只能画出其中一部分振荡

宇宙学家担心，人们会过于望文生义地将宇宙的“起点”等价于t=0的时刻。我们知道，当宇宙的年龄小于10^{-43}秒时，量子理论对整个宇宙的影响就变得非常强烈了。这个时间也正是“搅拌大师”宇宙最早开始物理上可区分的振荡的时刻。但相比振荡的频率，宇宙的膨胀速度太快了，以至于即使这些振荡都能一直持续到现在，也就能发生十几次。在现实的宇宙里，相比无穷多次，这十几次振荡太少了，这样光就无法在早期在很远的距离内反复传播。事实上，几乎所有的振荡都发生在 t=0之后的一瞬间，正如图 7.9 中那样。在靠近起点的无限小处还真是“繁忙”啊。

要想理解当前的宇宙结构，“搅拌大师”宇宙并不是关键所在，抛开这点儿小失望不说，它还是非常有意思的。它具有的行为是人们能从爱因斯坦方程组中找到的、最一般同时也是最复杂的行为。它是否会是爱因斯坦方程组所允许的最一般的宇宙呢？还不好下结论。

到了1980年，宇宙学家们曾经的信念，即关于宇宙甚早期的物理机制能够让宇宙变得各向同性而均匀的想法，再也行不通了。需要抹平的不规则性的类型太多了，而所有抹平过程都严重地受限于光速的大小，以及今天宇宙中的辐射所包含的熵。

磁性宇宙

考官：什么是电？

考生：哦，老师，我确定我学过什么是电（我以前肯定知道），但是我已经忘记了。

考官：多么不幸啊。只有两个人知道什么是电，一个是造物主，还有一个就是你。可是现在，其中一个人已经忘记了。

——牛津大学自然科学面试，大约在1890年[30]

对各向异性宇宙突如其来的兴趣，使得宇宙学家们开始思考，这样的宇宙中是否存在一种特有的能量类型。在此之前，他们研究了一些非常简单的例子，如黑体辐射和他们叫做“尘埃”的物质——由星系或恒星组成的没有压强的“气体”①（毕竟星系和恒星之间不常发生碰撞）。可是，宇宙中并不仅仅存在能量和粒子。磁场就是困扰人们已久的巨大谜团之一。

① 从整体上研究宇宙时，人们常常把星系和恒星看作均匀分布的、相对论压强为零的气体。在宇宙学语境中，有人把这种组成称为物质或尘埃（这个概念有时也包括暗物质）。与之并列的是正压强的辐射和负压强的暗能量。——译者注

磁场遍布整个宇宙。行星上、恒星上、星系中都有磁场。这些磁场是从哪儿来的呢？宇宙本身会不会就拥有一个大型磁场，所以最后才会产生这样的小型磁场？有一种可能性是说，这些磁场在宇宙诞生之初就产生了，同所有的物质和辐射一起，随着宇宙的膨胀而逐渐被削弱、变得稀薄。后来，它们被万有引力压缩进了众多星系的范围之内，又随着星系的自转而有所增强，产生了如今我们在恒星之间观测到的强大磁场。我们仍然不清楚，磁场的起源是确实就这样简单，还是说早期宇宙存在某种复杂的机制，在某一段时间里将这些磁场随机产生了出来。

如果宇宙刚刚诞生时就有了磁场，那么宇宙膨胀的速率就必须随着方向的不同而改变，这样才能维持压强和电压的各向异性，以使这个宇宙尺度的大型磁场能够继续存在。1965～1967年间，美国天体物理学家基普·索恩（Kip Thorne）、苏联宇宙学家雅克布·泽尔多维奇和安德烈·多罗什科维奇（Andrei Doroshkevich）发现了爱因斯坦方程组一类新的解。这种解描述了一种各向异性的宇宙，包含普通物质或辐射，并拥有一个宇宙尺度的大型磁场。他们证明，这种宇宙的开端必须也是各向异性的。但他们又发现，这种磁场在宇宙晚期的演化中又会产生很奇怪的效应。它会急剧地减缓各种各向异性在宇宙膨胀过程中的减弱过程，如

果它的强度足够大，就会在微波背景辐射中留下一个标志性的印记。但在实验观测中，微波背景辐射并不存在这样的印记，这使得宇宙学家们对于任何宇宙级大磁场的强度可以作出严格的限制。[31]

布兰斯–迪克的宇宙

不要相信任何实验结果，除非有理论作出过预言。

——无名氏

20世纪60年代，出现了种种质疑爱因斯坦广义相对论正确性的声音。天文观测表明，水星的运动轨道与爱因斯坦预言的竟然不一样。最终，人们在更加精确地了解了太阳表面的湍流活动（也就是说，太阳的直径和形状）之后，才解决了这个问题。但与此同时，这又引发了新的问题。罗伯特·迪克曾将地质学和古生物学的相关证据与天文学相结合，用新的方法检验引力理论的有效性。1966 年，他测量了太阳的形状，发现结果与广义相对论有矛盾。直到1973年，这些矛盾才被其他的观测结果解释清楚。[32]

让我们回想一下狄拉克的大数假设中关于牛顿引力“常数”G可能不是常数的想法。1961 年，迪克和他的研究生卡尔·布兰斯发展了一个出色的新理论，对爱因斯坦理论做了一般性的推广（图 7.10a）。[33]他们将万有引力常数写成了一种场的形式，就像物质的密度和温度一样，会随着时间和位置的变化而改变。如果把工作做到位的话，即不像狄拉克那样，只是将 G当作方程中的变量，可实际上 G又不在方程中变化，一个严格自洽的新理论就出现了。G不是任意变化的。它必须保证能量和动量的守恒，并且像其他形式的能量一样，能够弯曲周围的空间，改变时间的流逝。如果 G 退化成一个常数的话，那么布兰斯和迪克的理论就退化为爱因斯坦的理论。然而，如果允许 G 非常缓慢地变化的话，就可以解释水星运动的观测结果了（后来，人们发现这种解释其实是错误的）。

布兰斯和迪克的宇宙与爱因斯坦理论所预言的宇宙非常相似。有的永远在膨胀，有的在体积达到最大值后开始收缩，但其中略有差别。万有引力常数减小的速度可以比狄拉克原来的理论慢得多。如果G随时间的变化率是诸如$1/t^n$的形式的话，那么宇宙的尺寸就按照$t^{(2-n)/3}$的形式膨胀。所以当 n等于零的时候，G就不随时间变化，我们就得到了爱因斯坦和德希特的、大小按照$t^{2/3}$变化的简单宇宙。[34]

通过研究大量太阳系内的观测数据和引力实验的结果，人们发现，爱因斯坦理论预言的水星运动与观测精确地符合，而 n的取值会越来越趋近于零。到了1976年，人们对这种理论已经失去了兴趣。不过，二十五年以后，人们又萌生了新的兴趣，因为参照这种理论，就可以将爱因斯坦理论做出其他形式的推广，把另外一些传统的自然常数设为变量，并且可以自圆其说。

自1999年以来，人们通过观测类星体的光谱发现，越来越多的证据表明，在一百亿年以前，决定原子物理的一组自然常数可能和现在的略有不同（例如一个电子携带的电荷大小、光速或者电子质量）。[35]这些观测可以解释为，一个由不同自然常数组合而成的“精细结构常数”[36]会随着宇宙年龄的增加而略有增大，大约比宇宙膨胀的速率慢一百万倍。我们之所以要在1999年观测类星体的原因，也就是这些观测实验的新颖之处在于，比起实验室中的研究对象，类星体对精细结构常数的变化要敏感得多。

这些观测结果需要用一种推广的爱因斯坦引力理论才能解释，这种理论中能够自洽地引入变化的精细结构常数（支配着电和磁的强度）。2002 年，在雅克布·贝肯斯坦[37]的基础上，哈佛·桑德维奇、荣·马戈依卓和我[38]发现了一个类似于布兰斯和迪克的理论，并对精细结构常数随时间变化的行为作出了一个简单的预言（图7.9b）。在天文学尺度上，电和磁的效应并不重要，所以尽管宇宙膨胀的速率决定了这个“常数”随时间缓慢增大的速率，这个理论却对宇宙的膨胀没有任何可观测的影响。

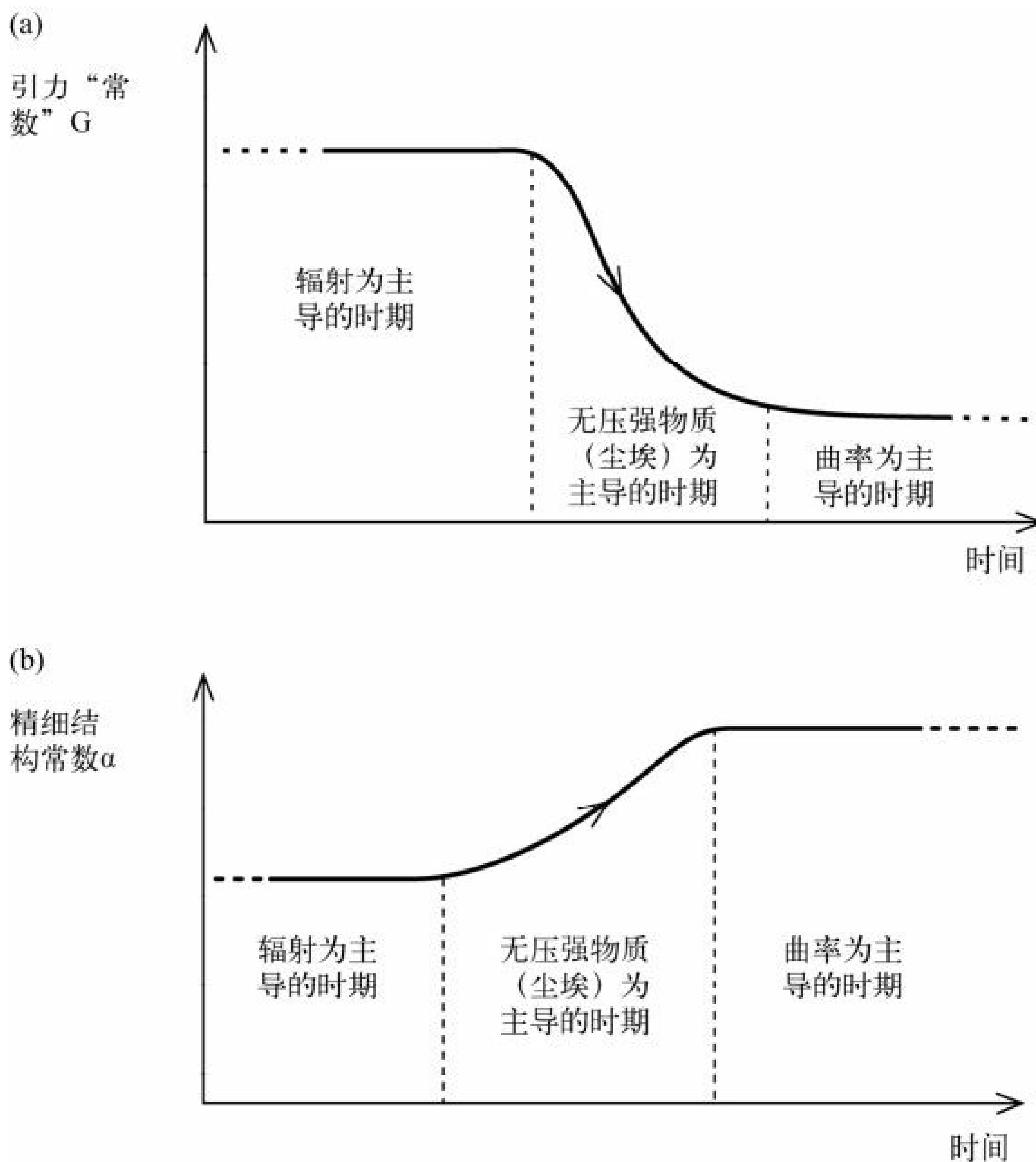


图7.10 (a) 布兰斯和迪克的引力理论预言，所谓的引力“常数”G实际上是随时间变化的，此图表示它典型的演化行为。G 的取值只有在无压强物质（“尘埃”）在宇宙中占多数时才会发生明显的变化。在宇宙以辐射为主导的时期或者在曲率为主导的开放宇宙中时，G是不变的。(b)该图表示桑德维奇、巴罗和马戈依卓的宇宙学理论所预言的精细结构常数α随时间的变化。在宇宙以辐射为主导或曲率为主导的时期，α不发生变化。在“尘埃”为主导的时期，α正比于宇宙年龄的对数

物质–反物质的宇宙

哎呀，能有什么事情呢？①

——英国童谣[39]

① 取自童谣“Oh Dear! What Can the Matter Be?”。童谣描述了小女孩望眼欲穿，期盼父亲从集市上及早带回礼物。作者利用了matter一词的另一个意义：物质会是什么样子？——译者注

1965年，微波背景辐射的发现，使得人们突然对宇宙的大爆炸理论产生了浓厚的兴趣，同时也抛弃了它的老牌竞争对手，例如稳态模型。关于我们身在“哪个宇宙”的争论，主要集中在炙热的早期宇宙究竟是各向同性的、各向异性的还是完全混沌的。但是也有一些人根本不买大爆炸理论的账。或者说，有这么一些人固守稳态模型在哲学意义上的优越性，因为这种宇宙没有开端也没有终点：它是可能存在的对称性最高的宇宙，因为它同时具有空间上和时间上的对称性。还有一些人既不看好大爆炸模型，也不看好稳态模型，而是寻求完全不同的理论，来描述膨胀宇宙甚早期时可能发生的事情。其中一个非常有趣的理论来自于反物质在宇宙中所扮演的角色。1928年，狄拉克预言了自然界中应该存在反物质。[40] 1932年，卡尔·安德森（Carl Anderson，1905～1991）从宇宙射线中第一次发现了反电子（又叫“正电子”）。后来，加州大学伯克利分校的欧文·张伯伦（Owen Chamberlain）、埃米利奥·塞格利（Emilio Segre）、克莱德·维刚德（Clyde Wiegand）和汤姆·伊普赛兰提斯（Tom Ypsilantis）在1955年发现了反质子。

没过多久，反物质就在宇宙学中引发了一场争论。20世纪60年代中期，两位顶级的瑞典物理学家汉纳·阿尔芬（Hannes Alfvén，1908～1995）和奥斯卡·克莱因（Oskar Klein，1894～1977）提出，宇宙的开端应该包含等量的物质和反物质，并处于一个体积巨大的低密度状态——这和宇宙大爆炸理论说的完全相反。物质和反物质在万有引力的作用下缓慢地相互靠近，最终，粒子和反粒子发生碰撞，继而湮灭变成海量的辐射。这场爆炸阻止了一些粒子和反粒子继续向中心掉落，使它们转为向外膨胀。这样的过程可能不是到处都有，不过阿尔芬和克莱因认为，这就是我们现在所观察到的宇宙膨胀，因为我们必然处于其中一个粒子–反粒子发生过湮灭的区域，由收缩转为了膨胀。

阿尔芬并不是一个宇宙学家，却是等离子体物理学和磁学领域的专家，他在这方面的工作赢得了1970年的诺贝尔物理学奖。阿尔芬的物质–反物质宇宙学在学术圈内几乎没引起什么注意，而且很快就有人发现这个模型与一些观测结果不符。[41]这个模型要求宇宙中物质的总和必须等于反物质的总和：有一个原子，就得有一个反物质原子；有一颗恒

星，就得有一颗反物质恒星。我们并没有发现任何反物质原子、反物质行星或者反物质恒星，于是正反物质湮灭时所生成的区域就必须格外广阔，以使物质和反物质完全分离——可是，这又谈何容易呢？更糟糕的是，宇宙从收缩转为膨胀时，其物质密度必须比我们现在观测到的小一百倍。宇宙中不太可能发生过这样一次反弹。

20世纪60年代末期，有一些在高能物理学方面更内行的科学家，比如罗兰德·奥姆斯（Roland Omnès）尝试在大爆炸理论的基础上，发展出一套物质–反物质对等的宇宙学理论。他的初始假设还是宇宙包含了相同数量的物质和反物质，并且这种对称的状态会一直保持下去；物质Ḿ反物质的比例不变，这在当时被认为是自然界牢不可破的守则。[42]为了解释我们周围的宇宙中没有任何反物质存在的证据，这个理论中必须存在纯物质和纯反物质构成的孤岛。

如果我们倒转时间的方向，逆转宇宙的膨胀过程，这些孤岛就会彼此相互靠近。一开始的时候，这些孤岛必然混作一团，形成一个充满辐射的热平衡态，其中粒子和反粒子不断在辐射的海洋中生成和湮灭。这发生在宇宙最初的千分之一秒内。我们可以计算出当等量的物质和反物质平均地混在一起，并停止湮灭时，会产生什么结果。对这个理论来说，结果是个坏消息，湮灭发生得非常充分，泽尔多维奇和丘宏义在1965年时证明，每产生10^{18}个光子，才能残留下一个质子或反质子。[43]而在我们现在所观测到的宇宙中，平均每个质子大约对应 10^{9}个光子。对于等量的物质–反物质的湮灭过程来说，我们的宇宙现在包含的质子和原子太多了，而产生的热辐射又太少了。

就这样，反物质的概念从宇宙学中销声匿迹了。但很快事情就发生了变化。20世纪70年代，宇宙学中发生了一场革命，反物质又被拉回了舞台中央，并引发了宇宙学研究史上最强烈的一次思想碰撞。

注释

[1] 这段话是1932年，他在伦敦的英国科学促进会上说的，转引自：M. Tabor,*Chaos and Integrability in Nonlinear Dynamics,* Wiley, New York (1989), p. 187. 1976年，海森堡临终时也说过类似的话。

[2] 如果宇宙的各个角落可以拥有不同的性质，那么就得用偏微分方程来描述，而不是常微分方程。这样的话，就大大增加了问题的难度，无论是借助人力还是计算机，都不会觉得轻松。

[3] 这个问题研究的是纳维尔–斯托克斯方程组（Navier-Stokes equations）。这组方程描述流体的运动规律，可以看作是著名的牛顿第二定律在流体力学中的对应版本。对于这个难题，更详细的介绍请见克雷数学研究所的网页：http://www.claymath.org/millennium.

[4] C. F. Von Weizsäcker,*Z. f. Astrophysik* 22, 319 (1943).

[5] C. F. Von Weizsäcker,*Naturwissenschaften* 35, 188 (1948).

[6] C. F. Von Weizsäcker,*Astrophys.* J. 114, 165 (1951).

[7] 1941年9月，海森堡去了哥本哈根，专程拜见了尼尔斯·玻尔。后来人们纷纷进行推理和猜测，想知道当时他们两人到底有没有就制造原子弹的可能性交换意见。迈克尔·弗雷恩（Michael Frayn）以此为主题创作了戏剧《哥本哈根》。历史学家大卫·C.卡西迪对此作了总结，参见：David C. Cassid, ‘A Historical Perspective on *Copenhagen*’,*in Physics Today*, July 2000, pp. 28–32; *Uncertainty: The Life and Science of Werner Heisenberg*, W. H. Freeman, New York (1992).情况似乎是，这次访问是由德国政府的文化宣传部资助的，海森伯和魏扎克都为公众做了有关太阳系起源于湍流问题的演讲。

[8] G. Gamow,*Phys. Rev*. 86, 231 (1952).

[9] 角动量守恒定律要求，如果一个漩涡的质量为 M、半径为 r、转动的速度为 v，那么Mvr就是恒定不变的。对于通常的物质来说，M是不变的，所以角速度v正比于1/r。在宇宙的早期阶段，辐射曾经占过主导地位。由于宇宙膨胀导致的引力红移效应，辐射的M会正比于1/r，所以此时的v是不变量。这个简单的原理同时也适用于均匀膨胀的宇宙中旋转扰动的演化行为，栗弗席兹在1946年首次用广义相对论研究了这个问题。

[10]这叫做“惯性范围”（inertial range）。

[11]如果转动的速度是v、时间是t而漩涡的大小是L的话，那么v=Lt且能量传递率正比于v^2/t；消去t，我们就会发现v^3正比于L。这叫做柯尔莫哥洛夫谱。

[12]我在牛津大学的博士学位论文中推导了这个问题，参见：J. D. Barrow, ‘The Synthesis of Light Elements in Turbulent Cosmologies’,*Mon. Not. Roy. Astron. Soc.* 178, 625 (1977).

[13] J. Binney, ‘Is the Flattening of Elliptical Galaxies Necessarily Due to Rotation?’,*Mon. Not. Roy. Astron. Soc.* 177, 19 (1976).

[14] G. Lemaître,*The Primeval Atom*, Van Nostrand, New York (1950).

[15] A. Taub, ‘Empty Spacetimes Admitting a Three-Parameter Group of Motions’, *Annals of Mathematics* 53, 472 (1951).

[16] L. Bianchi,*Memorie di Matematica e di Fisica della Societa Italiana delle Scienze, Serie Terza.* 11, 267 (1898).罗伯特·詹特森将这译成了英语，参见：http://www34.homepage.
villanova.edu/robert.jantzen/bianchi/#papers.

[17]当时的宇宙学家对各种宇宙模型的对称性和其他千姿百态的性质（以及复杂性）都作了研究，而这是首次系统地用群论的方法对宇宙

模型进行了分类。

[18] R. B. Partridge and D. T. Wilkinson,*Phys. Rev. Lett.* 18, 557 (1967).

[19] Jean de la Fontaine, ‘Le Juge arbitre fable XII, 28, 4’ (1693), in Marc Fumaroli, (ed.), La *Fontaine: Fables*, 2 vols. Pub. Imprimerie Nationale, Paris (1985); 或者参见：http://www.jdlf.com/lesfables/livrexii/.

[20] C. W. Misner, ‘Neutrino Viscosity and the Isotropy of Primordial Blackbody Radiation’,*Phys. Rev. Lett.* 19, 533 (1967).

[21]当然，这些宇宙学家的观点并不像戈尔德说的那样完全经不起推敲。他们要么认为现在还没有能力解决复杂的问题，要么猜测可能还存在其他尚未发现的物理原理，能够保证初始条件拥有高度的对称性。即使人们发现微波背景辐射以后，戈尔德还是死硬地支持稳态宇宙模型。他知道，在稳态宇宙中，由于物质是不断创造出来的，所以必然会导致现在的宇宙看起来是各向同性而均匀的。霍伊尔和他的研究生贾扬·纳利卡（Jayant Narlikar）曾经证明，稳态宇宙中的任何不对称性都会迅速消失，恢复到均匀膨胀的状态。主要问题在于，这种宇宙的平滑机制效率太高了，以至于人们根本无法解释在这样的宇宙中如何才能形成恒星和星系。

[22] J. D. Barrow and R. A. Matzner, ‘The Homogeneity and Isotropy of the Universe’,*Mon. Not. Roy. Astron. Soc.* 181, 719 (1977).

[23] http://blog.djmastercourse.com/harmonic-mixing-mixing-in-key/.

[24] C. W. Misner, ‘Mixmaster Universe’,*Phys. Rev. Lett.* 22, 1071 (1969).

[25] J. D. Barrow,Phys. *Rev. Lett.* 46, 963 (1981).

[26]这种搅拌器是由Sunbeam Products生产的，现在它是 Jarden公司的子公司。

[27]这种过程是混沌的。如果你将两个非常相似的“搅拌大师”宇宙放在一起比较，经过几轮振荡以后，这两个宇宙就会大相径庭。值得注意的是，描述混沌演化的机制却是确定论的（deterministic），这有些像一个无理数可以表示成一个无限的连分数。参见：V. A. Belinskii, E. M. Lifshitz and I. M. Khalatnikov, *Sov. Phys. Usp.* 13, 745 (1971).

[28]几何级数 1/2+1/4+1/8+1/16…，其中每项都是前一项的一半，把无穷多项都加起来就等于1。

[29] C. W. Misner,*Phys. Rev.* 186, 1328 (1969).

[30]转引自：Falconer Madan, in *Oxford outside the Guide-Books*, B. Blackwell, Oxford(1923).

[31] J. D. Barrow, P. G. Ferreira and J. Silk,*Phys. Rev. Lett.* 78, 3610

(1997).

[32] C. Will,*Was Einstein Right?*, Basic Books, New York (1993).

[33] C. Brans and R. H. Dicke,*Physical Review* 124, 925 (1961). 1955年，帕斯库尔·约尔丹曾发展出一个类似的理论，但这个理论是写在一本德文书《引力和宇宙》［*Schwerkraft und Weltall, Vieweg*, Braunschweig (1955)］里，可能是因为约尔丹在“二战”时期的所作所为不得人心，这本书当时没有引起什么关注。霍伊尔在BBC的演讲以及根据演讲汇集成的《宇宙的本质》［*The Nature of the Univers*, Blackwell, Oxford (1950)］一书中对这个理论嗤之以鼻。

[34] J. D. Barrow, ‘Time-Varying G’,*Mon. Not. Roy. Astron. Soc.*, 282, 1397 (1996).

[35] J. D. Barrow and J. K. Webb, ‘Inconstant Constants’, *Scientific American*, June 2005, pp. 56–63.更全面的介绍参见：J. D. Barrow,*The Constants of Nature*, Jonathan Cape, London (2002).

[36]精细结构常数是由电子的电荷 e、光速 c 和普朗克常数 h 所定义的无量纲量，即^{2}2 ✓e /hc。测量精细结构常数的实验达到了非常高的精度，这个数值大约是 1/137。它全面掌控了原子和分子的结构。

[37] J. D. Bekenstein,*Phys. Rev.* 25, 1527 (1982).

[38] H. Sandvik, J. D. Barrow and J. Magueijo,*Phys. Rev. Lett. 88*, 031302 (2002); J. Magueijo,*Faster Than Light*, Penguin Books, London (2003).

[39]《牛津童谣辞典》认为这首歌的歌词形成于1770～1780年间。

[40] “反物质”一词的历史要古老得多，可以追溯到物理学家阿瑟·舒斯特（ Arthur Schuster，1851～1934）的一篇理论性的文章：‘Potential Matter: A Holiday Dream’,*Nature* 58, 367 (1898)。他在文章中假设原子就像一种不断向宇宙输入能量的东西。不过该文中的反物质和狄拉克在1928年提出的反物质的严格定义没多大关系。

[41] G. Steigman,*Ann. Rev. Astron. Astrophys.* 14, 339 (1976).

[42]这叫做重子数守恒。重子数用来衡量物质粒子和反物质粒子（如质子、反质子）的数目差异。人们假设，无论粒子数和反粒子数怎么变，重子数都是不变的。

[43] Y. B. Zeldovich,*Advances Astron. Astrophys.* 3, 241 (1965); H. Y. Chiu,*Phys. Rev. Lett.* 17, 712 (1966).

第8章 混元之初

要是确实打算从零开始做苹果馅饼，那你得先把宇宙创造出来。

——卡尔·萨根

奇异的宇宙

奇异之事几乎总是意味着线索。越是平常、越是普通的罪案，越是难以理清。

——阿瑟·柯南·道尔

几乎所有我们提到过的宇宙模型都有一个显著的共同点。它们都“诞生”于过去的某个有限的时刻，而且当时宇宙的密度无穷大。正是这种怪异的诞生方式，曾遭到稳态模型创始人的强烈抵触，激发了他们寻求一种不同的理论，试图让宇宙不用经历这样的特殊时期。早在稳态模型发明之前，理查德·托尔曼就已经试着绕开这种必然的开端了，他提出了一种闭合的宇宙，膨胀和收缩的循环永不停歇，就像一个反弹的皮球。宇宙的每次反弹都是一个悬而未决的疑问，因为爱因斯坦的理论在密度和温度极高的时候不再适用。你当然不能指望当宇宙的体积降到零、密度变成无穷大的时候，一切定律都还原封不动。

从爱因斯坦方程组中解出第一批宇宙模型时，人们还存在一些不同的看法，他们大都在怀疑宇宙是否真的诞生于密度无穷大的状态。起初，爱因斯坦认为这不过是因为我们忽略了宇宙中物质的压强。他认为，如果把压强也考虑在内，当我们倒转时间的流逝，回溯宇宙的过去时，我们就会发现压强会大到足以令宇宙停止收缩，进而发生反弹，导致宇宙开始膨胀。这就像在压缩一个气球，把它的体积变得越来越小。压强会反抗这个过程，直到我们再也无法进一步压缩。

可惜，爱因斯坦的直觉在这里是错误的。在他的引力理论中，任何形式的能量都产生引力，包括压强。和我们的直觉刚好相反，压强并不会阻止宇宙塌缩到体积为零的一点；实际上，它加速了塌缩的过程，使得无穷压缩的时刻提早来临：压强产生的引力加剧了引力塌缩。

接下来，爱因斯坦设想，所谓的密度无穷大会不会是由于认为宇宙膨胀速率呈现了完美的对称性而人为得到的结果呢？如果你把球对称膨胀的过程倒过来，在过去的某个时刻，所有的一切都会汇聚到一个点上。但如果宇宙的膨胀并不完全符合球对称呢？如果你倒转膨胀的过程，宇宙的各部分就会彼此“错过”，这样就可以避免密度无穷大的状态了。在爱因斯坦看来，所谓的密度无穷大是一种“错觉”[1]，其他一流的宇宙学家也持此观点，例如罗伯特逊[2]和德希特[3]，那时还是在20世纪30年代早期。

这些简单的直觉都来自牛顿之类的旧体系，很快就被证明是错误的。1932年，勒梅特考虑了一种非球对称的、各向异性的宇宙。他证

明，这种宇宙也有一个密度无穷大的开端，就像各向同性的宇宙一样。[4]卡斯纳的各向异性宇宙和托尔曼的非均匀宇宙也都体现了同样的无穷大性质。[5]当然，可能还会存在其他更复杂的不对称性，例如旋转的宇宙可能就不会经历那么一个密度无穷大的状态。这个密度无穷大的状态，后来人们称之为原初“奇点”①。

① 此处，奇读作qí。奇点意为奇异的点。如果时空中包含了奇点（singularity），我们就说这种时空是奇异的（singular）。——译者注

无穷大的东西总是让人吃惊。如果在地球表面一些特定的地方散步，你就可能捡到 40亿年前形成的岩石；结构最原始的细菌大约出现在 30亿年以前，而现代人的祖先大约出现在20万年以前。地球和太阳系的年龄比这些表层岩石大不了多少，大约是46亿年。然而，宇宙的膨胀表明，如果我们向过去再追溯大约三倍的时间（138 亿年以前[6]），时间就不存在了，宇宙也不存在了，什么都不会存在了。这是一个惊人的结论。这时我们看上去非常接近万物的开端了。

人们认识到，压强和不对称性并不能简单地驱走原初奇点的幽灵。除此之外，在20世纪60年代早期，还有一种颇有影响力的看法认为，奇点并不是宇宙模型的物理性质，因此我们并不需要为奇点的物理实在性操心。受到朗道的启发，一组由尤金·栗弗席兹领导的苏联物理学家提出，大爆炸的奇点以及所谓的密度无穷大和时间的开端都是假的，不会产生危害。[7]借用一个熟悉的类比，设想地理学家有一个地球仪，上面覆盖了经线和纬线构成的网络，这样就能给地球表面上的每一个地点分配一组独一无二的标记。我们称之为“坐标系”，因为这能让我们标注地球上的不同地点。当我们朝着北极和南极看时，我们就会发现所有的经线都汇聚到了极点上：地图的坐标系退化到了一个特殊类型的“奇点”上。但这并不意味着地球表面上会有奇怪的事情发生。我们只不过选取了这样的坐标系来绘制地图。我们完全可以在极点附近换上另外一套网格，这样地图的坐标系就不会在那里失效了。

这就是栗弗席兹和他的同事所提出的观点，当我们倒转宇宙的膨胀回溯过去时，就会发生类似的事情。大爆炸只不过是由于描述宇宙时选取了不好的坐标系，从而导致了一个无害的地图奇点。[8]如果碰到这个假想的奇点时，我们就应该换一套表述，如果一套不行，那就再换一套。苏联的研究小组因而总结说，在物理学意义上，大爆炸的奇点并不是真实的，它并不是宇宙的开端。不幸的是，最终人们意识到，这种通过不断更换地图坐标系来避免奇点的办法仍然是一种错觉：它无法说明当你不断更换坐标系时到底发生了什么事情。人们在进一步研究后发

现，一个真正的物理奇点，就像地球表面上一个真正的洞一样，是不会通过坐标系的连续变换而被移除的。

20世纪60年代早期的时候，由此导致的困惑迫使宇宙学家们更加认真地思考奇点究竟代表了什么含义。如果把宇宙空间想象成一张随时间变化的床单，那么奇点就是床单上密度变成无穷大的点。现在，假设我们可以把这些点都剪下来扔掉。于是我们就得到了一张新的（带孔的）床单，能够描述一片完美无瑕的宇宙空间，不含奇点。这看起来有点作弊的嫌疑。从某种意义上讲，这样带孔的宇宙显然是奇异的。但是如果我们找到了一个无奇点的宇宙，我们如何知道用这种寻找和描述宇宙的方法，会不会像刚才剪床单那样，不恰当地将真正的奇点“剪掉”了呢？

1963年，米斯纳提出了一种回答，立刻改变了宇宙学家对奇点的理解。[9]这个想法抛弃了传统的奇点定义，不再把奇点看成是物质的密度或者其他什么物理量变成无穷大的那一点。我们说的奇点，应该是时空中任何物质的粒子或者光线的运动路径（包括它的一切“历史”路径）的终点，无论如何，路径也不会再延伸下去。

还有什么爱丽丝漫游奇境般的经历比沿着其中这样一条路径运动更为“怪异的”呢？在路径的终点，光线抵达了时空的边缘。再往前就什么都没有了。

用这样简单的方法定义奇点的妙处在于，如果一种物理性质，如物质的能量或密度在某个地方变成了无穷大，那么光线走到那儿就会停下来，因为那里的时间和空间都消失了。即使这个危险的地方被不恰当地从宇宙的地图上切除了的话，光线的历史也将在切除奇点后留下的洞的边缘处终结（图8.1）。

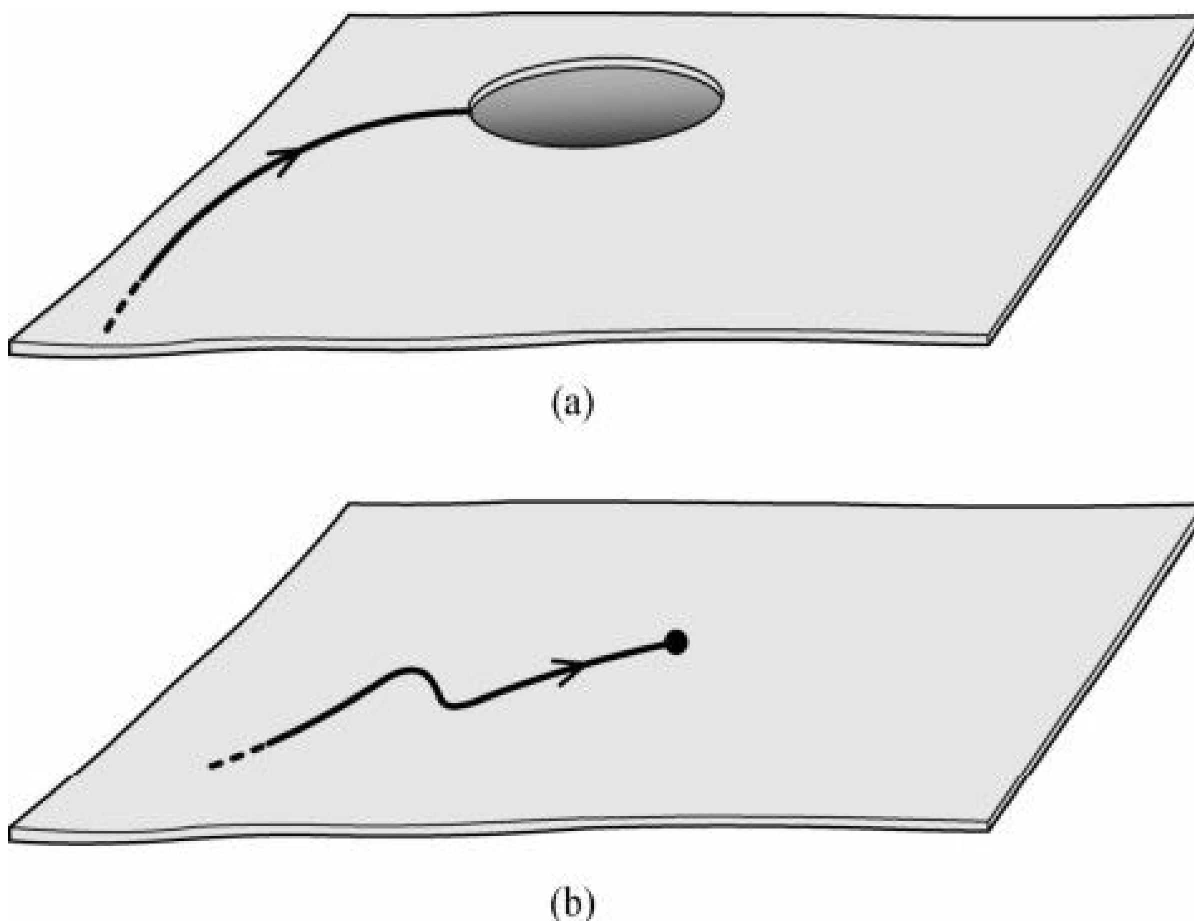


图8.1 这两个薄片表示光线会在宇宙中停止传播。(a)中有一个洞，光线止步于洞的边界；(b)中的光线跑到了一个时空被破坏了的奇点中

把奇点看成是时空边缘的方法有用极了。它绕过了爱因斯坦所担忧的关于宇宙对称性的问题，也避免了地图坐标变换所带来的模棱两可的问题。如果一个宇宙是非奇异的，那么其中任何粒子、任何光线的任何可能的历史都必须能够被无限地追溯：任何历史都不能有起点。一个非奇异的宇宙没有孔洞，没有边缘，也没有丢失的点。

这种想法简洁明了，不过也带来了不少疑难问题。可能的情况是，每次追溯到一段历史的起点时，那里的密度、能量或者温度都会是无穷大，从而导致了时空的消失，就像先前我们在弗里德曼–勒梅特宇宙中发现的直观想法，宇宙的开端是大爆炸。然而，我们并不清楚，一般来讲，这样的想法正不正确；直到今天，这个问题仍然没有标准答案，不过物理量的无穷大是众多膨胀宇宙模型的普遍性质。

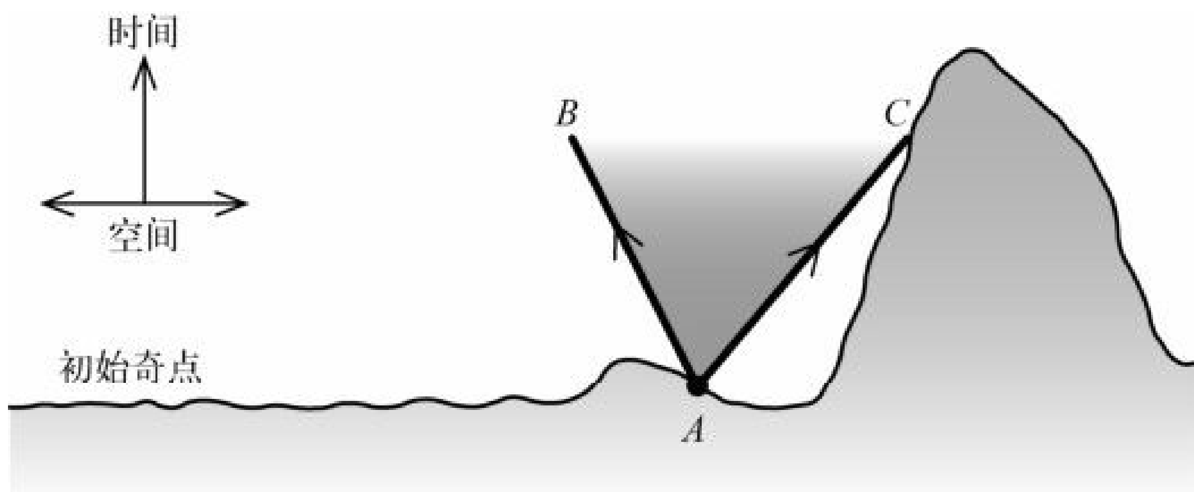


图8.2 初始奇点并不一定是同时出现的。宇宙某些地方的开端，如A点，是可以被其他地方的开端，如C点看到的，因为C点所代表的时刻比A点晚得多。线段AB和AC都代表从A发出的光

关于宇宙的开端，所要打破的常识不止以上这些，后面还有更加惊人的。奇点并不一定是普遍的，并不是所有的历史都需要有一个开端：其中一些历史从某一时刻开始，而另一些则可以向过去无限地追溯下去。而且即使所有可能的历史都有一个开端，这些历史也不一定是同时开始的。最具戏剧性的是，某些历史的开端可能会被其他地方的天文学家观测到（图8.2）。

哪些宇宙是奇异的呢？

奇点是多么奇怪的一种小东西啊，古怪的性质以及它并不存在的定义……这是我们某天必须加以理解的问题。

——罗伯特·格罗赫（Robert Geroch，美国理论物理学家）

在经历了定义奇点时的混乱和模糊，纠结于奇点究竟是否可以避免的问题之后，20世纪60年代早期，宇宙学家们想要找到一种明确的方法来研究一般宇宙中的奇点问题，而不是按照米斯纳的方法简单地研究每一个已知的宇宙解，一个接一个地看每一种宇宙有没有开端，有没有奇点。印度数学家阿马尔·库玛·瑞查杜里（1923～2005）[10]和美国的阿瑟·科马（1931～2011）[11]已经默默地于 1953年和 1956年分别开始着手这个问题了。他们都不喜欢奇点。可能是因为没有受到天主教传统观念的影响，瑞查杜里并不指望宇宙一定要有一个开端。即使过去的某个时刻里，密度变成了无穷大，他也不觉得这就意味着宇宙的开端。也许有某种未知的方法可以让我们穿过无穷大的地方，继续向过去追溯。他认为更可能的情况是，一旦密度变得过大，爱因斯坦方程组就不能再完备地描述宇宙了。爱因斯坦方程组中可能会出现新的项，方程组的解也可能

会发生变化，而这些变化了的解中可能就不存在奇点了。

这也代表了爱因斯坦对奇点的态度。就像两千多年前的亚里士多德一样，他相信宇宙中不应该存在物理的无穷大。如果宇宙的数学描述中出现任何形式的无穷大，要么是因为我们对宇宙模型进行了错误的简化，要么是模型所用的数学理论的假设失效了。

计算爱因斯坦方程组时，科马和瑞查杜里都发现，在一般的条件下，不对称且不旋转的宇宙仍然会在过去某个时刻处于密度无穷大的状态，正如弗里德曼和勒梅特的各向同性宇宙一样。

将旋转宇宙排除在外非常可惜，因为在牛顿的引力理论中，这本能防止宇宙进入密度无穷大的状态。但是很快，1965年，有人填补了这个空白，并且彻底解决了以前的地图坐标系变换导致的模棱两可的奇点问题，同时也开启了宇宙学研究的新篇章。

罗杰·彭罗斯是一名年轻的研究代数几何的理论数学家。在自己哥哥的朋友丹尼斯·夏马（Dennis Sciama，1926～1999）的影响下，彭罗斯开始考虑爱因斯坦的理论和宇宙的奇点问题。20世纪50年代，夏马曾经同霍伊尔、邦迪和戈尔德一起研究过稳态宇宙模型，不过1965年发现了微波背景辐射之后，他就立刻投向了大爆炸模型的怀抱。

彭罗斯将他熟悉的理论数学的方法引入了宇宙学的研究中，用一种全新的视角来理解奇点问题。他另辟蹊径，没有像科马和瑞查杜里那样，证明爱因斯坦方程组会导致过去某个时刻的密度无穷大，没有一个接一个地研究不同的宇宙解，也没有像栗弗席兹和他的合作者那样，引入不确定的近似。利用米斯纳提出的奇点定义，他将问题的关键集中在了描述粒子和光线运动轨迹的方程组上。他证明，在非常一般的条件下，这些运动轨迹中至少有一条需要一个起点，不管宇宙模型的其他细节究竟如何。

1965 年，彭罗斯革命性地证明了第一条“奇点定理”。这个定理说明，如果一颗垂死的大质量恒星在自身引力的作用下继续塌缩形成黑洞，黑洞的中心就会产生一个奇点。[12]在随后的 1965 年和 1966 年，史蒂芬·霍金、乔治·埃利斯和罗伯特·格罗赫仿照彭罗斯的方法，将奇点定理推广到了整个宇宙。[13]

当宇宙学家将以前的直观想法，也就是大爆炸时的物理量无穷大，同奇点是粒子轨迹的起点这样的数学描述联系起来时，一种非常微妙的情况就慢慢显现了出来。虽然他们可以像弗里德曼、勒梅特、卡斯纳和陶伯那样，将两者在最简单的宇宙模型中协调好，并证明运动轨迹无法延伸就是因为在物理的奇点附近时空遭到了破坏，但这个结论却并不适用于所有的宇宙模型。

数学方法的滚滚洪流最终产生了一个非常有说服力的结果。1970

年，霍金和彭罗斯共同给出了证明。[14]除了数学过程过于艰涩之外，这个证明的另一个引人注目之处在于，其中用到的假设在原则上都是可以被天文学观测检验的。他们证明，如果：

1. 空间和时间充分地平滑（于是我们就不必人为地引入轨迹的起点）；

2. 时间旅行是不可能的（这样的话，你就无法通过向未来运动而实际上回到过去，借此来绕过开端）；

3.宇宙中有足够的物质和辐射（于是我们就有足够强大的引力）；

4.不存在排斥性的万有引力；

5.爱因斯坦的广义相对论是正确的；

那么，时空中至少存在这样一条轨迹，它必须有一个起点。

注意这里的逻辑。这是一条定理而不是一个理论。如果这五条假设都成立的话，最终的结论也必然成立。如果其中一条或几条假设不成立[15]，并不意味着我们可以避免起点的存在：这只不过意味着定理在此时不成立，结果不能确定。事实上，当其中一些假设不成立时，我们既可以发现一些奇异的宇宙，也可以发现一些非奇异的、没有开端的宇宙。[16]

这五个条件的作用并不相同。人们认为前两条假设完全合理——对绝大多数人来说，允许时间旅行比宇宙有一个开端还要糟糕。

第三个条件能够被天文学家所检验。考虑到彭齐亚斯和威尔逊不久前发现的微波背景辐射产生的引力效应，这条假设被证明是正确的。

在当时，第四个条件被认为是正确的。它等价于要求宇宙中的物质密度ρ和压强p的关系在任何时候都满足这样一个条件：

$$\rho + 3p/c^2 > 0 \qquad (*)$$

其中，c 是光速。举个例子，组成宇宙微波背景辐射的黑体辐射或热辐射（也就是很久以前曾经占主导地位的辐射）都满足$\rho + 3p/c^2$的关系，又因为物质的密度ρ总是正的，所以对于辐射来说，条件（*）就总是满足的，它产生的万有引力也就总是吸引的（图8.3）。从他们证明这个定理的时候起，直到1977年，人们都相信条件（*）对所有可能形式的物质都成立。没有任何理由能够让人怀疑这一点，除非你是那种用怀疑的眼光看待一切的人。

第五个条件的情况略有不同，很可能不成立。但是怎样才不成立？在哪儿不成立？牛顿的引力理论已经统治了两百多年。最终，我们知道，它被爱因斯坦的广义相对论取代了。但这两个理论在很大程度上是

重叠的。当引力很弱、物体的运动速度远远低于光速时，爱因斯坦的理论在许多方面都会越来越像牛顿的理论。当运动速度接近光速、引力强大到可以导致物体以光速运动时[17]，爱因斯坦的理论仍然成立，并且是自洽的，而牛顿的理论就失效了。

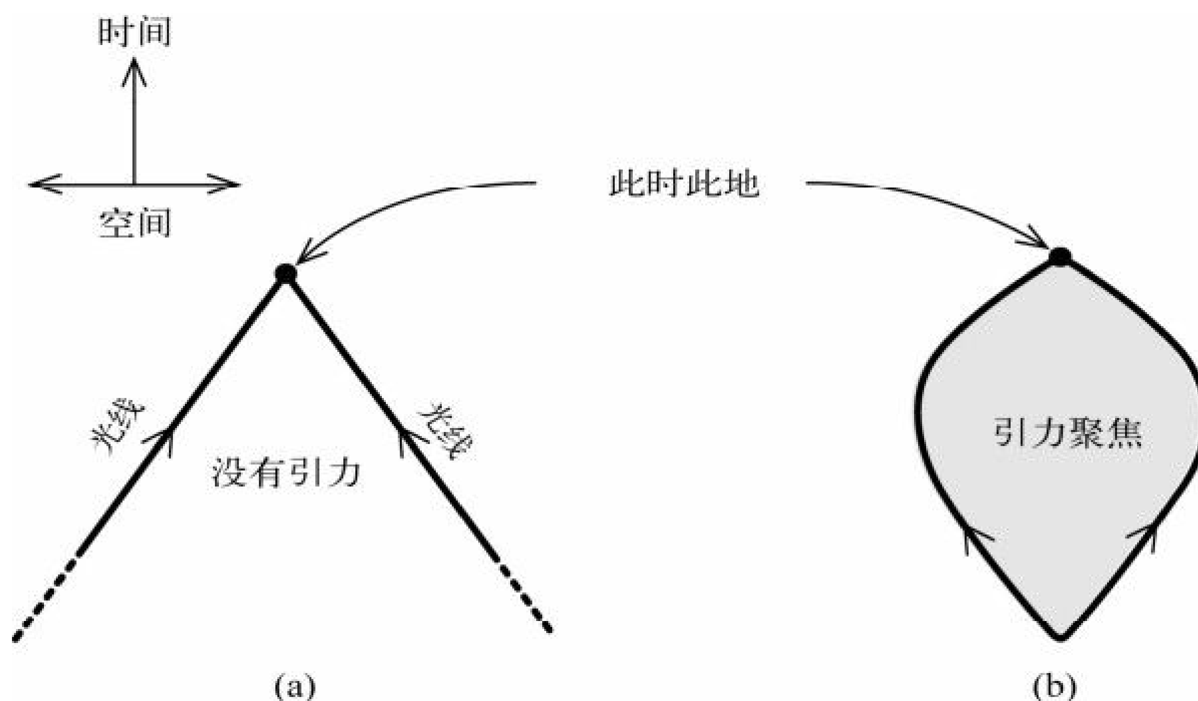


图8.3 (a)没有引力时，光线在时空中匀速传播的路径。(b)物质和辐射产生的引力使得光线传播的路径弯曲了。如果引力总是起吸引作用，宇宙中的物质又足够多的话，光线传播的路径在过去某一时刻就会聚焦在一个奇点上

考虑各种不同的极端情况时，我们估计爱因斯坦的理论最终也会失效。在20世纪的前25年中，玻尔、爱因斯坦、普朗克和海森堡发现，物质和能量的本质是量子的，能量的取值是特定的、一份一份的，并不是随便怎么取值都可以。传统物理学中的粒子都应该伴有波动的一面。相比水波，这种波动的特性更像是一股犯罪狂潮。这是一种信息的波动，正如一股犯罪狂潮波及你的街坊邻居时，周围很可能会出现犯罪行为。当电子的波经过你的仪器时，就意味着你很可能探测到一个电子。对一个质量为m的粒子来说，它的量子波长和m呈反比。所以粒子的质量越大，量子的波动就越小，如果这种波的势力范围比粒子本身的尺寸还要小得多，我们就不必太在意这些波的影响。然而，当一个粒子的质量非常小时，量子的波长就变得大了许多，能够轻而易举地超过了粒子的尺

寸。在这种情况下，以牛顿力学的标准来看，这种粒子的行为就看起来十分怪异：它的本质是量子的，也是波动的。

爱因斯坦的引力理论一点儿都没有将这种量子的怪异性考虑在内。正如约翰·惠勒在1957年所指出的那样，宇宙非常早期的时候可能存在这种量子的条件，爱因斯坦的理论在那段时期就会失效。此时，我们需要一个考虑了量子行为的新版本的广义相对论（“量子引力”）来接替广义相对论。为了搞清楚什么时候需要借助量子引力，我们要用一种简单的方法来判断，那就是考虑宇宙开始膨胀后光所走过的距离（“可见宇宙”的大小①），然后计算以这段距离为半径的球体中包含物质的总质量（当时的可见宇宙的总质量），看看这段距离在什么时候会小于这些物质的量子波长。在那个时刻之前，整个宇宙都会呈现出一种量子波动的行为，用爱因斯坦的理论是描述不了的。这段量子引力的时期发生在宇宙最初的10^{-43}秒之内，那时连空间和时间也变得不确定了。

① 宇宙一开始是不透明的，因为当时的光子与其他粒子发生剧烈的相互作用而被束缚住了。宇宙诞生后大约38万年时变得透明，光子终于可以自由地穿行了。假设那个时候你睁开了眼睛。开始的一刹那，你几乎什么都没看清。过了1秒钟，你已经能看到半径30万公里内的东西了（如不考虑空间本身的膨胀）。又过1秒钟，你的视线已经延伸到半径60万公里的区域里了。这就是你的过去光锥，也就是你的“可见宇宙”范围。那时到现在已经过了137亿年，我们的“可见宇宙”范围变得非常广袤，但仍然是有限的。在那个范围之外，仍然有一部分现在的我们所看不到的宇宙。由于宇宙正在加速膨胀，也许有一些地方我们永远也看不到。——译者注

设想我们逆着时间线追溯宇宙的膨胀，来到奇点定理预言的历史起点，并假设宇宙各处的历史是同时开始的。为了方便起见，我们将过去的这个时刻记为t=0的时刻。我们知道，当我们非常接近t=0的时刻时，不能指望爱因斯坦的理论能够完备地描述所发生的事情。我们估计在$t=10^{-43}$秒时，这个理论就失效了——这个时期叫做普朗克时期，以物理学家马克斯·普朗克（1858～1947）[18]命名。理论的失效导致任何有关的定理也不适用了，或者是因为空间和时间不再是光滑的，或者是因为量子引力的效应，导致霍金和彭罗斯的定理的第四个条件不成立了。引力可能变得具有排斥性，导致宇宙“反弹”，变回膨胀的状态，而不是继续收缩，密度越变越大。另外，我们还可以设想其他的可能（图8.4）。所有这些情况都符合宇宙晚期的膨胀行为。我们无法区分哪个才是正确的。

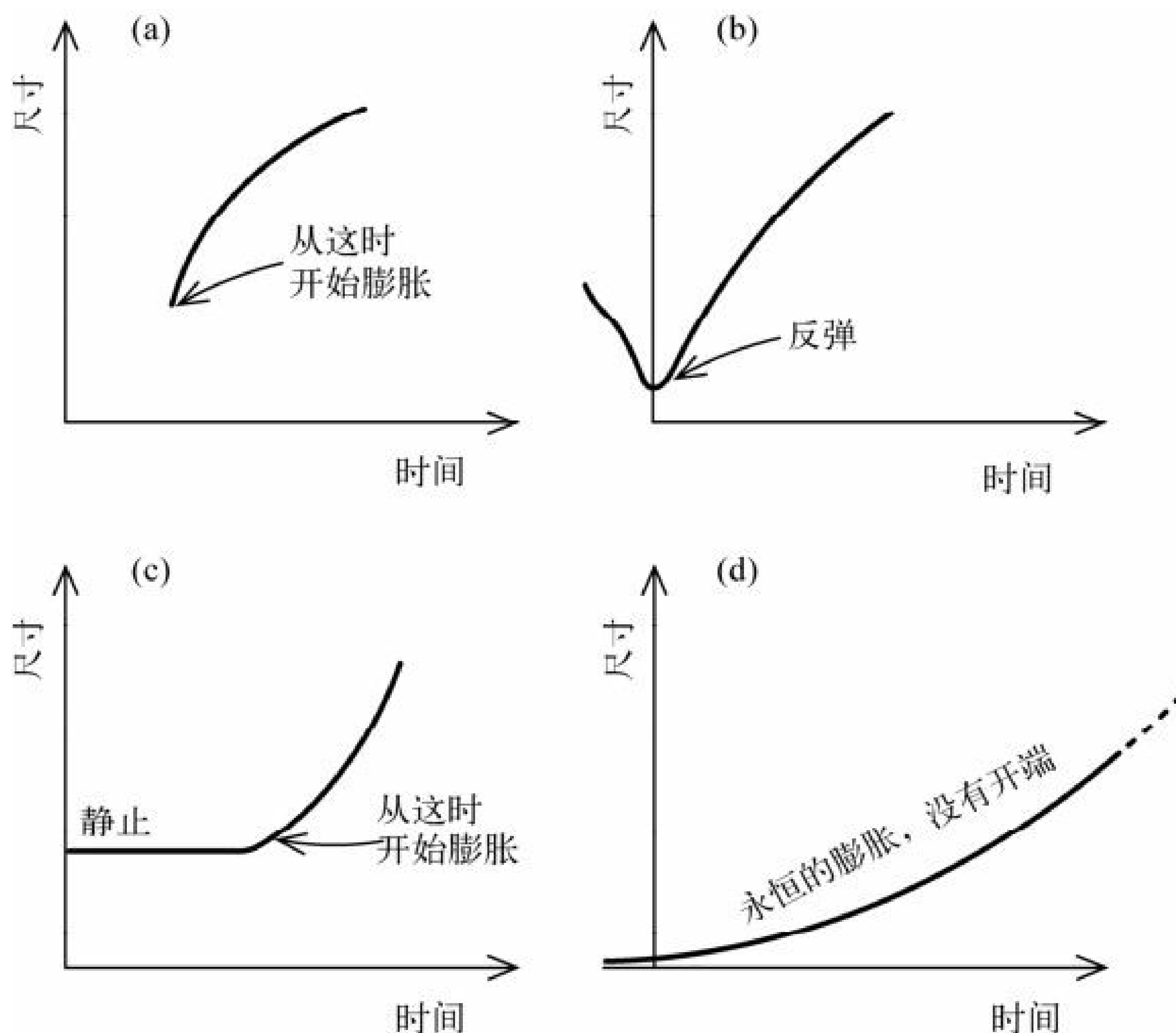


图8.4 一些理论假设宇宙的过去并不存在奇点，而是像今天所看到的那样在好端端地膨胀：(a)宇宙从过去某一时刻开始膨胀，但并不存在密度无穷大的奇点；(b)宇宙先是收缩，然后在密度有限大的某个时期发生“反弹”，接着开始膨胀；(c)在无穷远的过去，宇宙是静止的，然后逐渐开始膨胀；(d)宇宙一直在沿着一条指数曲线膨胀，在过去的任何一个有限的时刻，这条曲线的函数值都不为零

这让人想起理查德·托尔曼在1930年提出的旧的循环宇宙模型。奇点定理告诉我们，宇宙不可能一口气收缩成一个奇点，然后再反弹。因为在奇点处，空间和时间应该已经消失了：你不可能穿过奇点，而对爱因斯坦理论的失效视而不见。反弹必须在宇宙的尺寸很小，但没有变成零的时候发生，同时这又要求我们必须违反奇点定理的第四个条件。

尽管广义相对论在奇点处很可能已经失效了，但在 1966～1972 年间，宇宙学家还是在用奇点定理说明，宇宙早期必然存在一个密度和温

度都高得不可思议的时期。尽管爱因斯坦理论的失效避免了宇宙大爆炸时密度变成无穷大，但人们仍然认为宇宙的密度曾经高达水的密度的10^{94}倍。任何丰富的想象力都无法企及如此怪异的事情。

到目前为止，我们讨论的奇点都存在于过去，但如果我们生活在某个弗里德曼提出的“闭合”宇宙中，膨胀最终会转为收缩，奇点还是会在未来等着我们。当宇宙就要变成最后的奇点时，它必然会变得极为不规则，因为不规则性会随着时间的推移而越来越强大。黑洞会一个接一个地形成。宇宙的某些区域会率先经历大塌缩。一些观测者在经历大塌缩之前，甚至能观测到其他区域的观测者先走一步。不同区域先后发生的大塌缩会是一切的终点吗（图8.5）？这是什么意思？“终点”意味着什么？“一切”又意味着什么？并不是只有研究宇宙的起点时才会遇到如此尴尬、如此难以回答的问题。

过去奇点的存在引发了一系列元科学（meta-scientific）问题：在此“之前”发生了什么？宇宙从奇点诞生时，什么东西决定宇宙未来的演化？如果在奇点之前，空间和时间都不存在，那我们又如何证明物理定律在奇点之前也存在呢？你怎么能将通常的科学方法应用在像奇点这样独一无二的事物上呢？

这就是研究宇宙的起源时，人们所提出的一些怪问题。宇宙学家们把注意力集中在了一些有望回答的具体问题上。我们能否证明，当粒子的历史存在一个起点时，这个起点总是对应密度和温度的无穷大状态，就像我们逆着时间研究膨胀宇宙的模型时所遇到的情况？这个问题取决于我们是否认为奇点定理从本质上讲就是正确的。如果是，那么我们必须相信过去存在一个奇点。如果不是，那我们就全盘皆输。

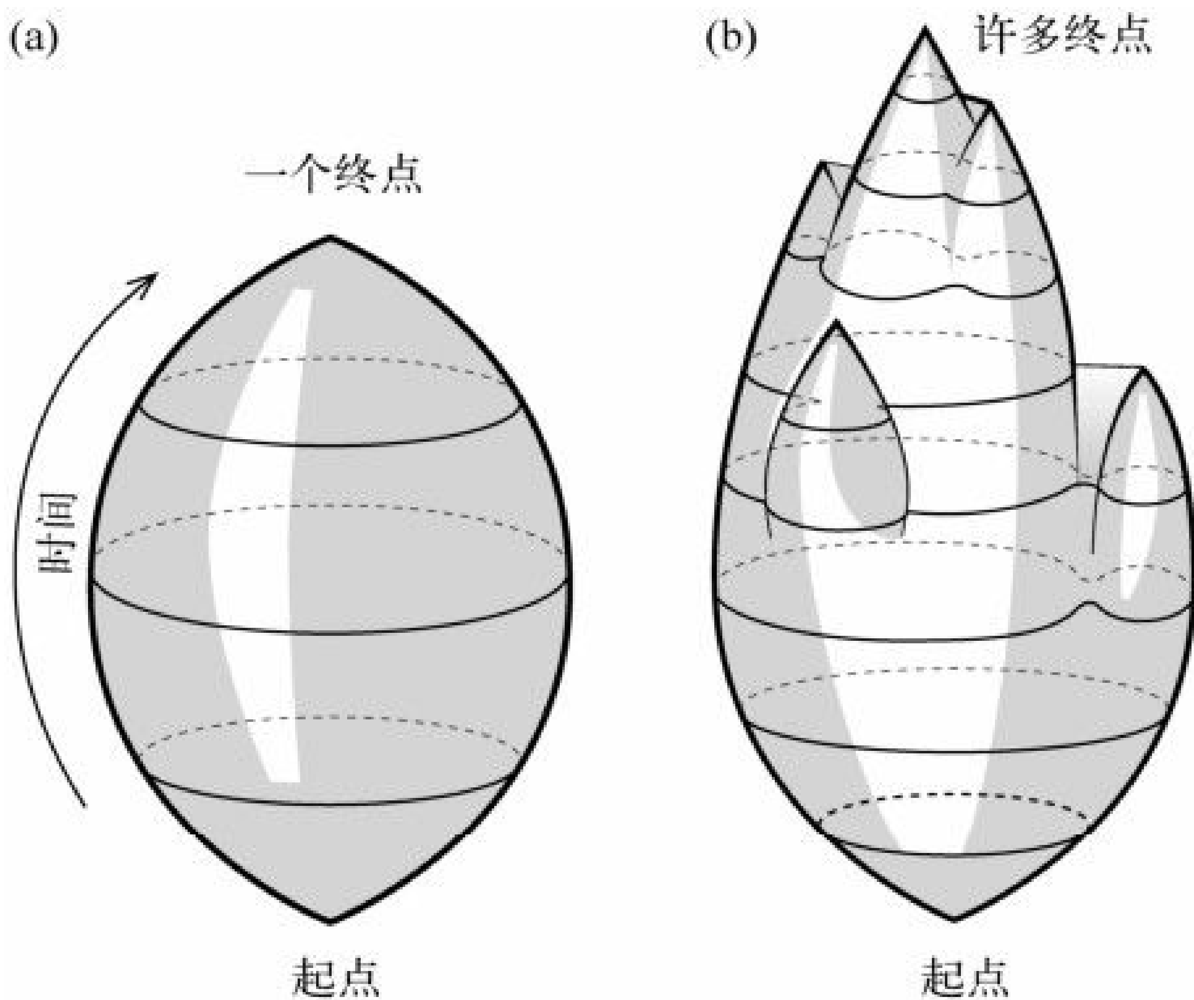


图8.5 (a)只有在一个绝对光滑的宇宙中，最后的奇点才会同时出现；(b)在真实的宇宙中，有些地方密度大而有些地方密度小，密度大的地方会更早地变成奇点。如果宇宙中有的地方率先变成了奇点，其他地方就有可能观察到这个现象

冷冰的宇宙和温热的宇宙

我知道你的行为，你也不冷也不热。我巴不得你或冷或热。你既如温水，也不冷也不热，所以我必从我口中把你吐出去。

——《新约·启示录》[19]

大爆炸理论的支持者长期以来坚持认为，宇宙曾经处于一个极为高温致密的状态。20世纪70年代末，这个观念又引出了一个新的宇宙学问题。关于宇宙的外形、宇宙如何变得光滑和各向同性、宇宙一开始到底是混乱不堪还是秩序井然的，在这些问题上我们已经强调得足够多了。但是我们并没怎么考虑过，宇宙到底是由什么构成的。微波背景辐射的发现导致了一个新的自然常数的产生：宇宙中光子数和质子数（或原子

数）的比值。这个数目非常大，大约是十亿。[20]

如果我们将宇宙中所有的物质搅匀，不再允许任何恒星和行星的存在，那么就剩下一堆均匀分布的单个原子，一立方米的空间中大约只存在一个原子。这个密度极其低，比地球上任何实验室中的人造真空都稀薄。在相同的体积中，大约存在十亿个微波背景辐射的光子。所以十亿比一的比例大体上体现了宇宙中热辐射物质的含量。十亿的数目太大了，比恒星爆炸或者今天宇宙中发生的其他剧烈过程所产生的光子数目还要大得多，这也是为什么我们难以用稳态宇宙模型解释背景辐射的存在的原因之一。当然，大爆炸理论没有解释，为什么光子数是原子数的十亿倍，为什么不是另一个超级大或者超级小的数目。

人们很快意识到，在热大爆炸宇宙的历史中，这个数目扮演了一个关键的角色。它决定了温度何时降低到可以形成原子，然后形成恒星和星系，以及这些天体所能形成多大的尺寸。但是为什么光子对原子的比例是十亿比一？这仍然是个谜。也有人尝试研究过，如果宇宙大爆炸之初是“冷冰的”，只有几个光子对一个原子，或者是“温热的”，一万个光子对一个原子，会不会有爆炸性的事件将这个数目自然而然地提高到十亿比一呢？[21]可惜，这些尝试并不成功。就像先前的湍流宇宙那样，这样的想法无法解释当前所观察到的氦元素的丰度（它们产生的氦元素太多了）。为了产生热辐射，所有可能的剧烈过程都会在微波背景辐射中留下明显的痕迹，使其不均匀度超过千分之一，而这是天文学家所观察到的上限。热的宇宙一开始就得这么热。

意外简单许多

在全部五幕戏中，他扮演的李尔王时刻处于恐惧中，就像有人要谋权篡位。

——尤金·菲尔德（Eugene Field，1850～1895，美国作家）[22]

这些研究产生了一个副产品，那就是寻找更多的方程，以描述大爆炸理论中宇宙早期的基本粒子的运动规律。这个问题首先由勒梅特提出，但最终未能解决。后来，阿尔珀、赫尔曼同詹姆斯·弗林一起在这个问题上迈出了第一步，将我们今天熟悉的高温物理学引入了大爆炸宇宙学。[23]

你可能这样想过，宇宙学家就是需要和核物理学家、粒子物理学家组成一个团队，这样才能了解宇宙年轻时应该长什么样。唉，可惜在1973年以前，高能物理学家可能还真帮不上什么忙。道理很简单，当时并没有理论能够有效地描述物质在一千亿度高温时的行为，这个温度远比地球上的实验中所能得到的任何温度都要高。理论家们也强不到哪儿去。他们试图为具有强烈相互作用的基本粒子建立一套理论，但所有的努力最终只是得出，能量越高，粒子间的相互作用就越强。很快，你就会面临一团乱麻，根本无法作出靠谱的计算。[24]这看起来就像为甚早期宇宙研究鸣起的丧钟。我们越是逆着时间追溯过去，粒子的行为就注定

越来越复杂，越来越棘手。所谓宇宙的“起点”，或者霍金和彭罗斯的奇点定理，似乎已经超越了我们所能理解的自然法则的范围。

1973年，在戴维·波利策（David Politzer）、戴维·格罗斯（David Gross）、弗兰克·维尔切克（Frank Wilczek）、徐一鸿（Tony Zee）和杰拉德·特胡夫特（Gerard ’t Hooft）的努力下，一个描述基本粒子高能行为的突破性理论诞生了。这个理论颠覆了高能基本粒子相互作用的旧图景。周围无处不在的量子真空能量海，会对遨游其中的基本粒子的性质产生一定影响。因此，当环境的温度增加时，粒子间的相互作用强度也会变化。将一个带电粒子，如电子，置入量子真空之中，就会从真空中吸引一些不断产生、又转瞬即逝的带正电的粒子，这些粒子的能量是从量子真空中借来的。于是，一个带负电的粒子就戴上了一个正电荷组成的屏蔽墙。当另一个能量很低的电子慢慢接近它时，如图8.6a所示，并不会受到电子所带的全部负电荷的影响，而只是感受其中未被和谐的部分。另一方面，一个高能电子，如图8.6b所示，会更加深入地穿透屏蔽墙，感受到更多未被和谐的裸电荷的影响。

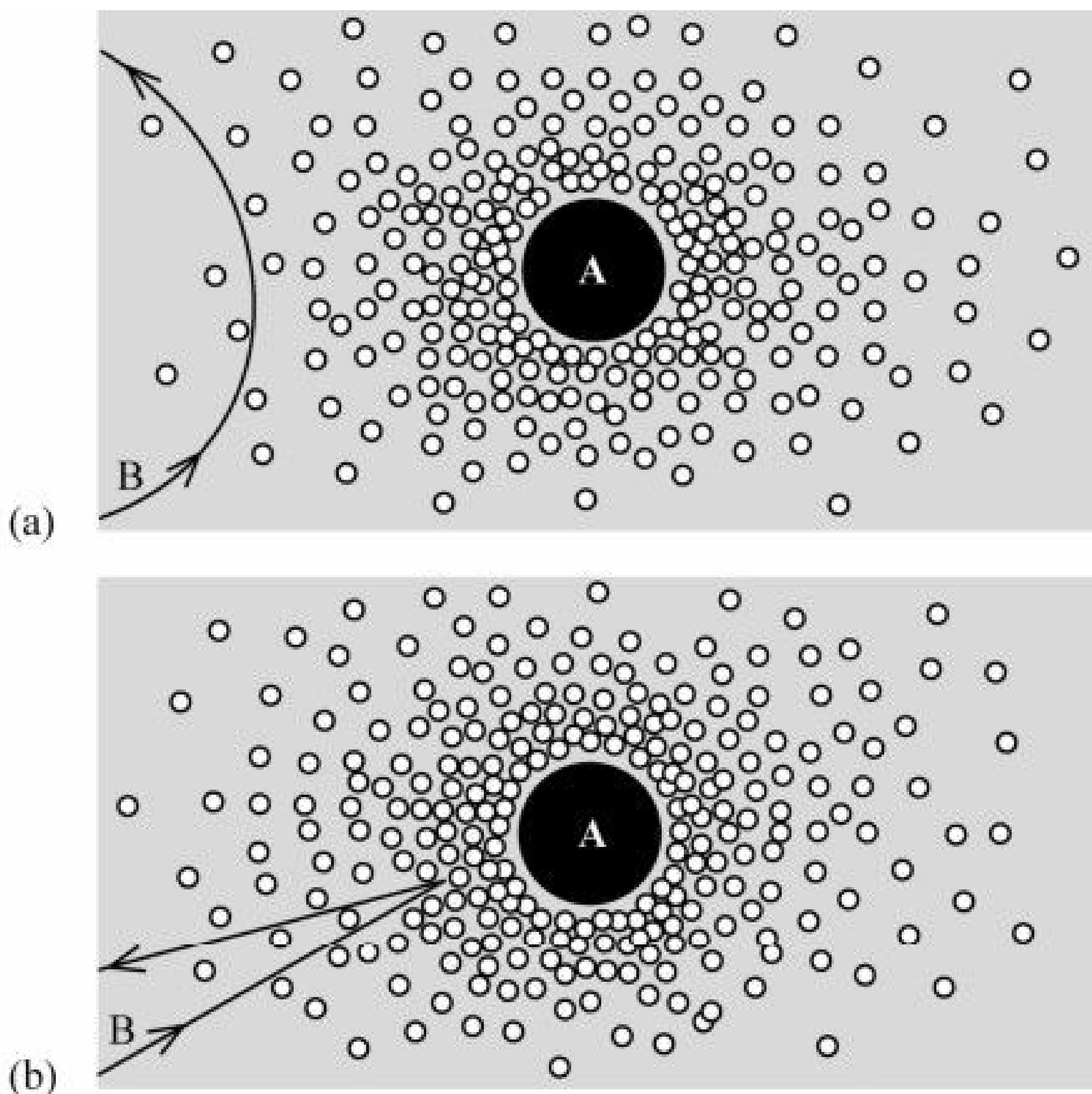


图8.6 (a) 一个入射的电子B，能量较低，而A的中心处是另一个带负电的目标电子，被周围的虚拟正电荷所屏蔽，因此两者相互作用较弱。(b) 一个高能的入射电子，能够穿透正电荷组成的屏蔽墙，因而能感受到目标电子所产生的未被和谐的更强的排斥力①

① 实际上 A和 B周围都存在虚拟的正电荷屏蔽墙，这里是用图像化的手段表示量子场论的核心内容，真实的物理机制要比图中的更完备。——译者注

这就意味着，相互作用的强度确实取决于周围环境的能量大小。对于参与电磁相互作用的粒子来说，例如电子，由于量子真空的存在，它

们感受到的等效相互作用强度会随着环境能量的增加而增强。对于参与强相互作用的粒子来说，例如夸克和胶子，量子真空产生的影响就完全在意料之外了：随着环境能量的增加，量子真空使得强相互作用的强度变弱了。这个获得了诺贝尔奖的发现叫做“渐进自由”，因为它表示说，能量最高的粒子之间就像没有什么相互作用，自由自在无拘无束。这也帮助宇宙学家们解开了旧的粒子物理理论的枷锁。突然之间，早期宇宙变得简单多了。

一统江湖

一就是一，仅此唯一，永应如此。

——《绿草如茵》[25]

似乎没过几年，这些理论的重大发现就改变了粒子物理学的方向。能量和温度增加时，粒子之间相互作用的等效强度会发生变化，这就为老谜题提供了新的解决方案：如果自然界中的相互作用强度各不相同，怎么会存在一种理论能将所有相互作用统一起来呢？表面上看，强相互作用和弱相互作用完全不同。不仅强度不同，连参与这两种相互作用的粒子的类型也不同。

电磁相互作用的强度较弱，如果弱的相互作用增强，同时强的相互作用减弱，就可能在很高的能量下，使亚核粒子之间的电磁、弱和强相互作用①的等效强度大小相等，三线交于一点。这种描述了三种基本相互作用的“大统一理论”（缩写为 GUT）的想法诞生于 1975年，当时霍华德·乔治和谢尔顿·格拉肖第一次找到了一个最简单的候选理论（图8.7）。[26]

① 强相互作用和弱相互作用分别是自然界两种基本相互作用的名称。强的相互作用和弱的相互作用分别指某种相互作用的强度有大有小。——译者注

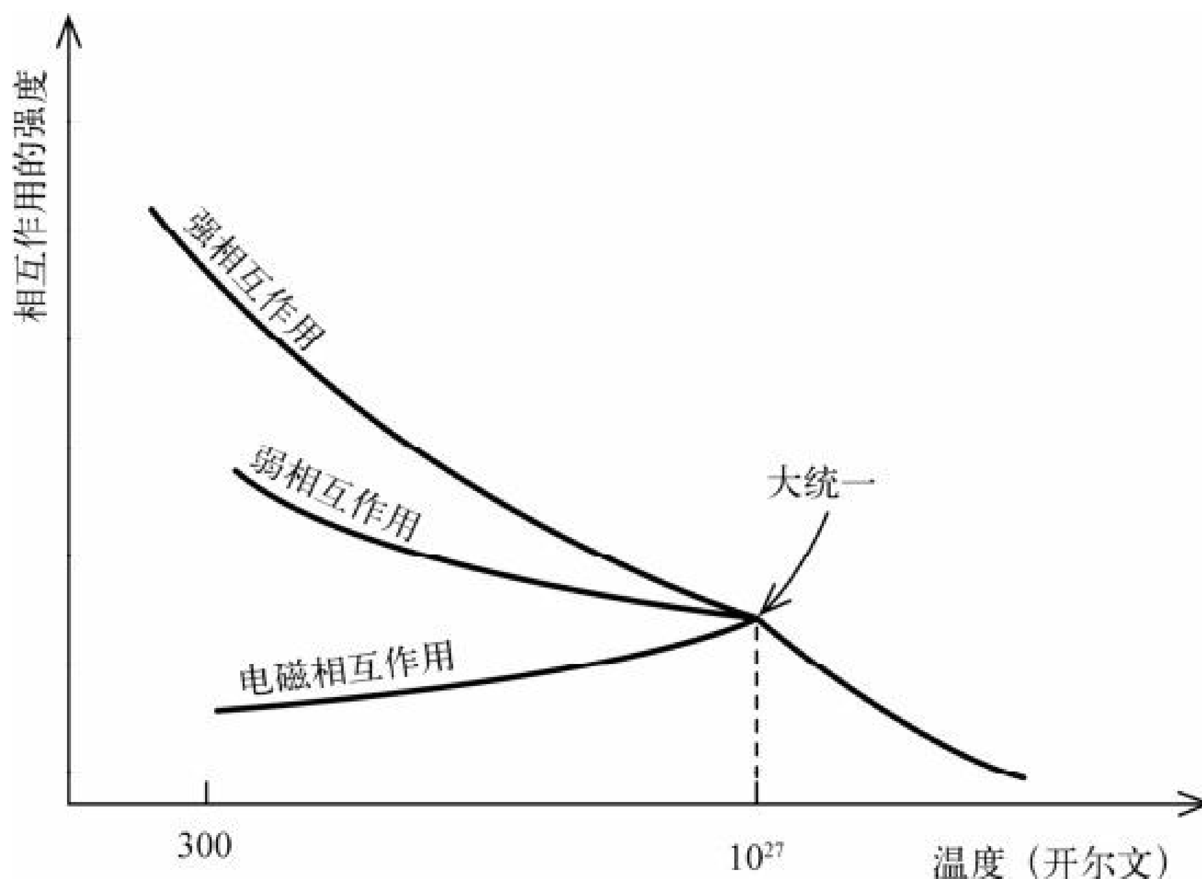


图 8.7 如果考虑真空涨落产生的效应，那么自然界中的强、弱和电磁相互作用的等效强度就会在非常高的能量下趋于一致，那时一个自然界三种相互作用的“大统一理论”就诞生了。现在我们知道，在低能的情况下，这三种相互作用的强度大不相同

这一切改变了宇宙学家们的观念，使得宇宙学的研究朝着另一个方向进行。突然之间，我们重构宇宙历史时所遇到的问题，例如如何描述宇宙甚早期时物质和能量的演化行为，不再是困难重重令人绝望，也不再注定越追溯越糟糕了。渐进自由意味着，当我们追溯宇宙膨胀开始时应该出现的那个高温时期时，相互作用强度就会越来越弱，事情也会变得越来越简单（至少没有变难）。所以宇宙学研究也发生了相应的变化。从前，人们都在一个劲地强调宇宙的几何性质，研究不同种类的膨胀行为，检验这些模型是否会演化成今天我们所看到的宇宙，并尝试探究宇宙是如何从一个过去的奇点中诞生的。粒子物理学家有时会路过打个酱油，加入到这些问题的讨论中，但没几个人会动真格儿。一旦渐进自由简化了物质的高温行为，使相关研究变得可行，主流粒子物理学家们便蜂拥到宇宙学中来。宇宙早期的环境为他们检验自己的理论提供了

便利的场所，他们可以在其中寻找相关的观测证据。

宇宙学研究的是大尺度上的物理机制，基本粒子物理学研究的是微小尺度上的物理机制，这两种物理学成功地融合到了一起。粒子物理学家想要检验自己的新理论，看看它们和天文学观测符合得好不好。宇宙学家开始研究高能物理学中粒子的反常性质，看看能否获得启发，帮助宇宙学解决一些伤脑筋的问题。太空中的不发光物质会不会是由一些未知的基本粒子构成的呢？为什么宇宙是由物质组成的，而不是由反物质组成的？

宇宙早期高温物质的研究透出一缕曙光，又把这样的问题重新照亮了起来。于是，一些物理量，如宇宙中物质和反物质之间的平衡，就不再是永恒不变的常数了。在这缕曙光的照耀下，新的思想和新的方法又会催生大量新的宇宙模型，必将使粒子物理学家和天文学家都叹为观止。

注释

[1] A. Einstein, *Sitz. Preuss. Akad. der Wiss.* (Berlin), pp. 235–237 (1931).

[2] H. Robertson,*Science* 76, 221–226 (1932).

[3] W. de Sitter,*Mon. Not. R. Astron. Soc.* 93, 628–634 (1933).

[4] G. Lemaître,*Publ. Lab d’Astronomie et de Géodésie de l’Université de Louvain* 9, 171–205(1932).

[5] 实际上，卡斯纳的宇宙并没有密度的无穷大，因为其中不包含任何物质。不过宇宙的膨胀速率和引力产生的潮汐力都会变成无穷大。即使你向卡斯纳的宇宙中引入物质，也不会经历密度无穷大的时期，就像包含无压强物质的海克曼Ḣ舒金（Heckmann-Schücking）宇宙模型一样。

[6] 目前对宇宙年龄的最佳估计是137亿±1亿年。

[7] E. M. Lifshitz and I. M. Khalatnikov,*Sov. Phys. JETP* 12, 108–113 (1961), 558–563 (1961), *and Advances in Physics* 12, 185–249 (1963); E. Lifshitz, V. V. Sudakov and I. M. Khalatnikov,*Sov. Phys. JETP* 13, 1298–1303 (1961).

[8] 有人（错误地）认为，虽然各向同性的宇宙包含奇点，但是各向同性的宇宙不够典型，因此最典型的宇宙应该不包含奇点。

[9] C. W. Misner, *J. Math. Phys.* 4, 924–937 (1963).

[10] A. Raychaudhuri, *Physical Review* 98, 1123–1126 (1955), *Physical Review* 106, 172–173(1957). 瑞查杜里的第一篇论文是在1953年4月份向杂志投稿的，结果一直拖到1955年才发表。

[11] A. Komar,*Physical Review* 104, 544 (1956).

[12] R. Penrose,Phys. *Rev. Lett.* 14, 57 (1965).

[13]对此研究的介绍参见：G. F. R. Ellis and S. W. Hawking, *The Large Scale Structure of Space-Time*, Cambridge University Press, Cambridge (1973).

[14] S. W. Hawking and R. Penrose,*Proc. Roy. Soc. London A* 314, 529 (1970).

[15]第二条假设“不存在时间旅行”，如有必要，也可以用另外一种假设替换。

[16]用数学的行话来说，这五条假设是过去存在奇点的充分（非必要）条件。

[17]此时引力势能的大小接近了光速的平方c^2。有意思的是，爱因斯坦的理论似乎对自身施加了某种限制：能够产生的最强大的力是$c^4/4G$，大概是3.25×10^{43}牛顿，或10^{39}吨。

[18]关于这段历史的来龙去脉，参见我的书［*The Constants of Nature*, Jonathan Cape, London(2002)］中的第2章和第3章。

[19] Revelation 3: 15–16.

[20]在发现微波背景辐射以前，曾经有人研究过“冰冷的”宇宙，这样宇宙的早期并没有以热辐射为主导的时期。苏联的泽尔多维奇［*Advances in Astronomy and Astrophys.* 3, 241 (1965)］和美国哈佛大学的大卫·雷泽［*Ann. Rev. Astron. Astrophys.* 3, 341 (1964)］都对这个模型作过简单的研究。1965年，彭齐亚斯和威尔逊发现了微波背景辐射以后，泽尔多维奇迅速领会了其重要性，不再坚持冰冷的宇宙模型。但直到1984年，雷泽认为微波背景辐射只不过是恒星发出的光被加热的结果，所以他仍然坚持宇宙早期十分寒冷，参见：D. Layzer,*Constructing the Universe*, W. H. Freeman, San Francisco (1984), chapter 8.

[21] M. J. Rees,*Phys. Rev. Lett.* 28, 1669 (1972); Y. B. Zeldovich,*Mon. Not. Roy. Astron. Soc.* 160, 1P (1972); J. D. Barrow, *Nature* 267, 117 (1977); B. J. Carr and M. J. Rees, *Astron. Astrophys.* 61, 705 1977).

[22]这是他对克莱斯顿·克拉克（Creston Clarke）扮演的李尔王的评价。

[23] R. Alpher, J. Follin and R. Herman,*Physical Review* 92, 1347 (1953).

[24]物理学家都喜欢这样的理论，在其中计算的过程一步比一步精确，计算结果所需要的改动一步比一步细微。而文中提到的计算过程却恰好相反——这一定标志着理论计算中含有某种严重的错误。

[25]作者不明，中世纪代代相传的民歌，内容综合了基督教、非基督教和天文学的内容。

[26] H. Georgi and S. Glashow,*Phys. Rev. Lett.* 32, 438 (1974).

第9章
美丽新世界

想象出来的宇宙比这个粗制滥造的“真实”宇宙漂亮多了。

——G. H.哈代（G. H. Hardy，1877～1947，英国数学家）

正反物质不对称的宇宙

大量时事评论员已经向这个问题倾注了大量愚昧无知之辞。如果他们继续这么搞，我们可能很快就会对这个问题一无所知了。

——马克·吐温（1835～1910）

20世纪 70年代后期，在粒子物理学家进入宇宙学研究领域后，人们开始运用新的大统一理论去解释宇宙中正反物质不对称的问题，以及为什么光子和质子的比例是十亿比一的问题。这些理论表明，自然界中三种非引力的相互作用虽然从强度上看有天壤之别，但还是有可能统一在同一个框架内的。随着环境温度的增加，这些相互作用的等效强度也会以不同的速率和方式发生相应的变化——弱的相互作用越来越强，而强的相互作用越来越弱。于是当我们追溯到宇宙的最初10^{-35}秒时，不同相互作用的强度在极高的能量下就殊途同归了。

这种殊途同归也为其他问题的“统一”提供了一种解决方案。现有的理论将每一种自然界的相互作用分开研究，它们不允许粒子之间进行任意的相互转换。夸克不能变成电子①，反之亦然。这像是误将基本粒子世界划分成了独立的小团体，不同团体之间连话都不说。小团体的划分依据是，粒子是否拥有特定的属性，例如电荷或“色荷”。后者是一种更复杂的属性，只有参与强相互作用的粒子才有。大统一理论预言了一类新粒子的存在，这类粒子同时携带两种荷，并且它们能在前面提到的各种粒子小团体之间扮演中间人的角色。这些中间粒子可能非常重，当环境的能量使得几种相互作用的等效强度变得一样时，这些粒子就会大量产生。它们被称为“X”粒子，在宇宙的早期阶段，它们在不同类型的基本粒子中间创造了一种真正的平等。

① 夸克参与所有四种基本相互作用；电子只参与三种，不参与强相互作用。因此在那些理论中，这是两类不可相互转化的粒子。——译者注

基本粒子之间的新型相互作用立即导致了两个后果，引起了物理学家的注意。如果带色荷的粒子，如夸克，会转化成像电子、中微子之类的粒子的话，那么质子之中的三个夸克也可能衰变。质子就应该是一种不稳定的粒子。

乍看之下，质子的平均寿命，大约是 10^{30}年，似乎不算长。只要屏蔽了宇宙射线和其他外界影响，我们就有可能在地下深层实验室中放置

的数千吨物质中观察到质子的衰变。尽管单个质子的平均寿命比宇宙的年龄（1.4×10^{10}年）还大得多，但质量为一吨的岩石或水中就含有大约10^{30}个质子。如果将这个质量放大一千倍，再在周围布满灵敏的探测器，我们就有很大的机会观察到质子衰变。[1]

印度科拉尔金矿地下实验室自1980年起投入运行后，有人声称从中观察到了质子衰变。于是人们群情激昂，但后来发现这不过是一次误报。[2]实验结果逐渐限制了质子平均寿命的可能取值，这要比先前的乐观估计大得多（大于6.6×10^{33}年[3]）。尽管在任何大统一理论中，质子都是不稳定的，但质子的寿命极有可能特别长，使我们无法从宇宙射线产生的鱼目混珠的“假衰变”事件中检测出真正的质子衰变来。[4]

新型相互作用导致的第二个后果是，夸克和它们的反粒子，也就是反夸克，不会以相同的比率衰变。这个性质结合夸克会衰变成电子和中微子的性质之后，就会导致宇宙的正反物质平衡发生变化。于是，我们有可能将这个平衡算出来，而且结果也不需要依赖于宇宙开始时的初始数值。

这个平衡取决于三个因素[5]：宇宙中能参与夸克–反夸克衰变反应的粒子所占的比例、夸克和反夸克衰变速率的差异，以及衰变速率的大小。之所以要考虑最后一个因素，是因为衰变速率必须比宇宙膨胀的速率快，这才能阻止逆反应的发生，防止物质–反物质的平衡再回到初始的大小。

为了计算这个平衡，人们做了很多尝试，细节越来越多，模型越来越复杂。[6]这些计算都得出了一个相同的结论：无论宇宙的初始状态如何，我们都确实有可能解释如今正反物质不对称的事实。此外，如果夸克和反夸克之间的差异约为十亿又一个质子对十亿个反质子的话，那么当宇宙冷却到10^{13}度时，每十亿个质子都会跟十亿个反质子湮灭成二十亿个光子，只剩下一个质子。因此，我们最终就得到了一个光子和质子的比例为二十亿比一的宇宙[7]，跟当前的观测非常接近。物质–反物质的平衡问题与光子原子的比例问题紧密地联系在了一起，可以被同时解释清楚。

我们不必过高地评价这些计算的重要性，并不是必须运用某种特定的大统一理论，或者某种机制来解释宇宙中正反物质的不对称性：这些计算表明，许多理论都能轻易地得出这种解释。当我们找到那个正确的大统一理论时，它很可能预言了正反物质的不对称和光子质子数目的比例。这些数字之所以现在如此，不再是因为它们过去就是那样。

问题宇宙

再也没有欣喜和充满信心的晨曦！①

——罗伯特·布朗宁（1812～1889）[8]

① 译文来自“天涯诗会”，作者为肥仔深。——译者注

没过多久，一个深奥的宇宙学问题就露出了丑恶的嘴脸，威胁着物理学大统一理论的安危。[9]如果宇宙刚诞生不久时，弱、强和电磁相互作用在极高的能量下变成同一种相互作用，那么就必然会产生一个不必要的副产品。电磁相互作用在早期宇宙中出现时，就会大量产生一种非常重的新粒子，叫做磁单极子，这种粒子首先是由狄拉克在1931年时提出的。[10]它就像一个非常重的电子（比电子重10^{20}倍），只不过不携带电荷，而是携带一种新型的“磁荷”，就像条形磁铁的一个磁极。这些自然界中的磁荷是守恒的，就像电荷一样有正有负，除掉磁单极子的唯一方法就是让正反磁单极子碰在一起，发生湮灭。可惜，一旦形成，磁单极子就不太有机会碰到反磁单极子，于是就在宇宙中保留了下来，而且数量非常多。

这就完蛋了。今天的宇宙应该早已被磁单极子占领了。这些磁单极子为宇宙贡献的能量密度是所有恒星和星系的10^{26}倍。在这样的宇宙中，不会存在恒星和星系，也就不会有“我们”在为此操心了。这就是“磁单极子问题”。[11]

将高能物理学嫁接到重构宇宙甚早期历史的尝试，在这个拦路虎的干扰下停住了脚步。[12]说来奇怪，破天荒头一回，不是新的天文观测结果，也不是爱因斯坦方程组的新解，而是粒子物理学为宇宙学研究的发展指明了方向。

暴胀的宇宙

可以大致看到未来事物演变的雏形。

——莎士比亚[13]

磁单极子的问题造成了某种僵局，但这些问题又吸引了人们的注意力，使他们主要将宇宙看作广袤的粒子物理学实验室，而不是看作恒星和星系所处的膨胀背景。粒子物理学家只对最简单的膨胀宇宙模型感兴趣：一个宇宙拥有平滑的各向同性的空间，没有转动，也没有其他不同寻常的属性。[14]

1980年，突然之间，一份新的打印稿出现在物理系的邮箱里。这篇文章内容充实，出自斯坦福直线加速器中心的一位年轻粒子物理学家之手。当时美国的学术市场非常低迷，他还没什么名气，已经做了好几期博士后这样的临时工作。[15]阿兰·古斯发现，有一种宇宙模型可以同时解决困扰了宇宙学家二十年的好几个问题。不过他的最初动机并不在此，他只是想要解决磁单极子的问题。

物理学家和宇宙学家很快发现，古斯想法的实质非常简单。它假设，在宇宙的甚早期，恰好在磁单极子的出现和正反物质的平衡建立起来之前，宇宙经历了一次短暂而突然的加速，或者按照古斯的叫法，是

一次暴胀（图9.1）——这是为了讽刺当时害得人找不到稳定工作的经济大环境[①]。

① 通货膨胀和暴胀的英文都是inflation。——译者注

这个想法可不简单。我们已经见过一些加速膨胀的宇宙，例如德希特宇宙和稳态宇宙模型。这些宇宙永远在加速——从过去到永远。但是我们也见过像勒梅特那样的宇宙，一开始是减速膨胀的，当宇宙学常数的排斥作用超过普通的引力时，这种宇宙就转而开始加速膨胀。一旦开始加速，就永远不会再停下来。

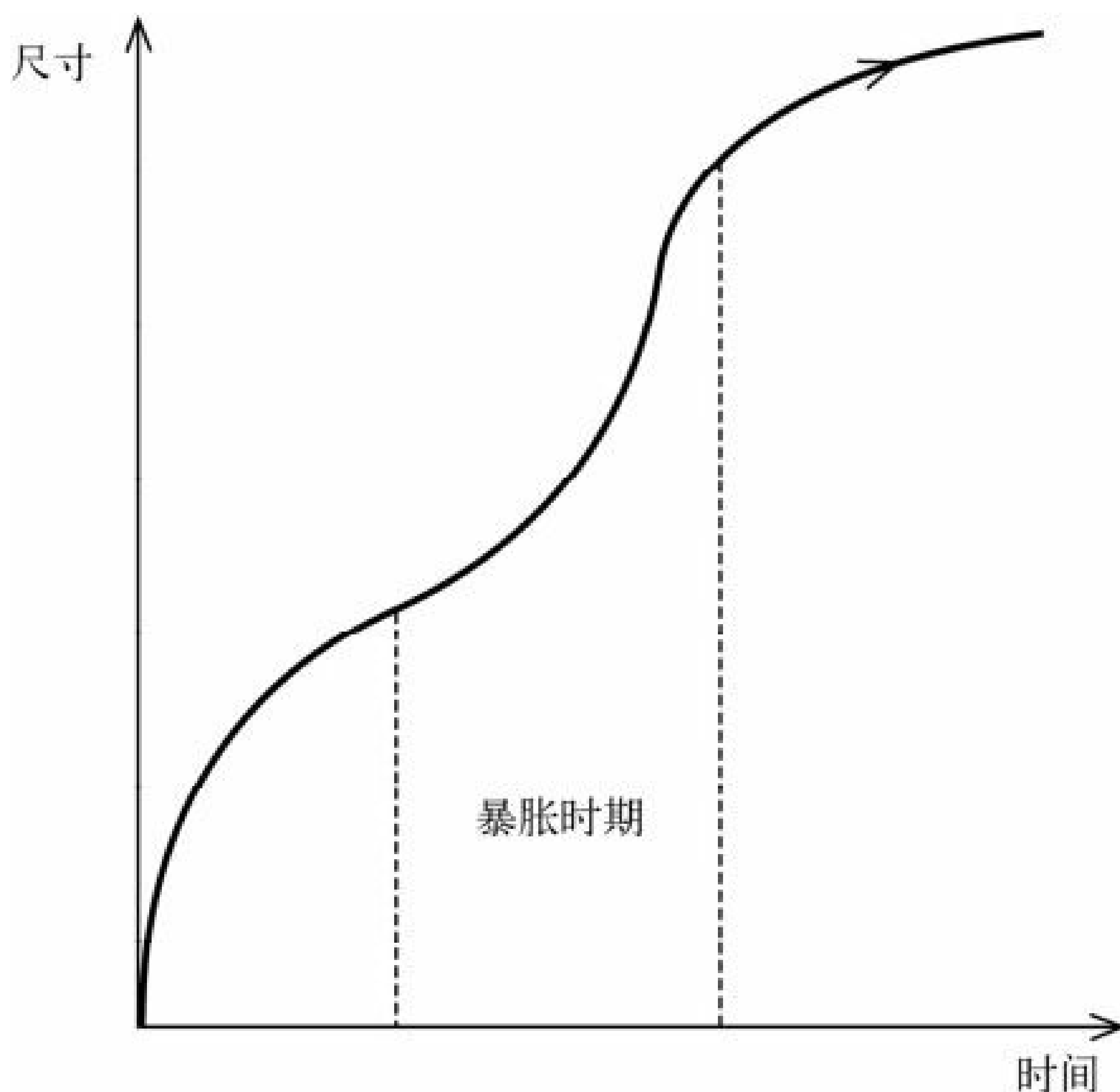


图9.1 一个简单的暴胀宇宙模型中，尺寸随时间变化的图。暴胀是说在一段有限的时间内宇宙发生了加速膨胀。暴胀导致宇宙膨胀得更遥远，也使得后来的膨胀速率更接近临界值

没有人见过加速以后还能再次减速的膨胀宇宙模型。古斯找到了一种办法，定义了一种暂时起排斥作用的引力①，在一段很短的时期内，这种排斥力就像爱因斯坦那个臭名昭著的宇宙学常数一样。

① 这种力仍然属于引力相互作用，只不过表现形式与万有引力相反，是相互排斥的。——译者注

这个新理论能够统一地描述物质的高温行为，其中包含了一种新型粒子，后来被称为“标量场”。[16]能量的这种形式能够缓慢地变化，比宇宙膨胀的速度慢得多，这样它们就会产生排斥性的引力，而不是相互吸引。在这种极端而又最常见的情况下，这种标量场简直就是爱因斯坦宇宙学常数的翻版。这种标量场的压强是负的，但和正宗的宇宙学常数的区别在于，它们的生命如白驹过隙：这种标量场或快或慢会发生衰变，转化成普通的辐射物质或基本粒子，于是压强变成正的，排斥力变成引力。图9.2是暴胀宇宙的温度和尺寸的总体演化图。

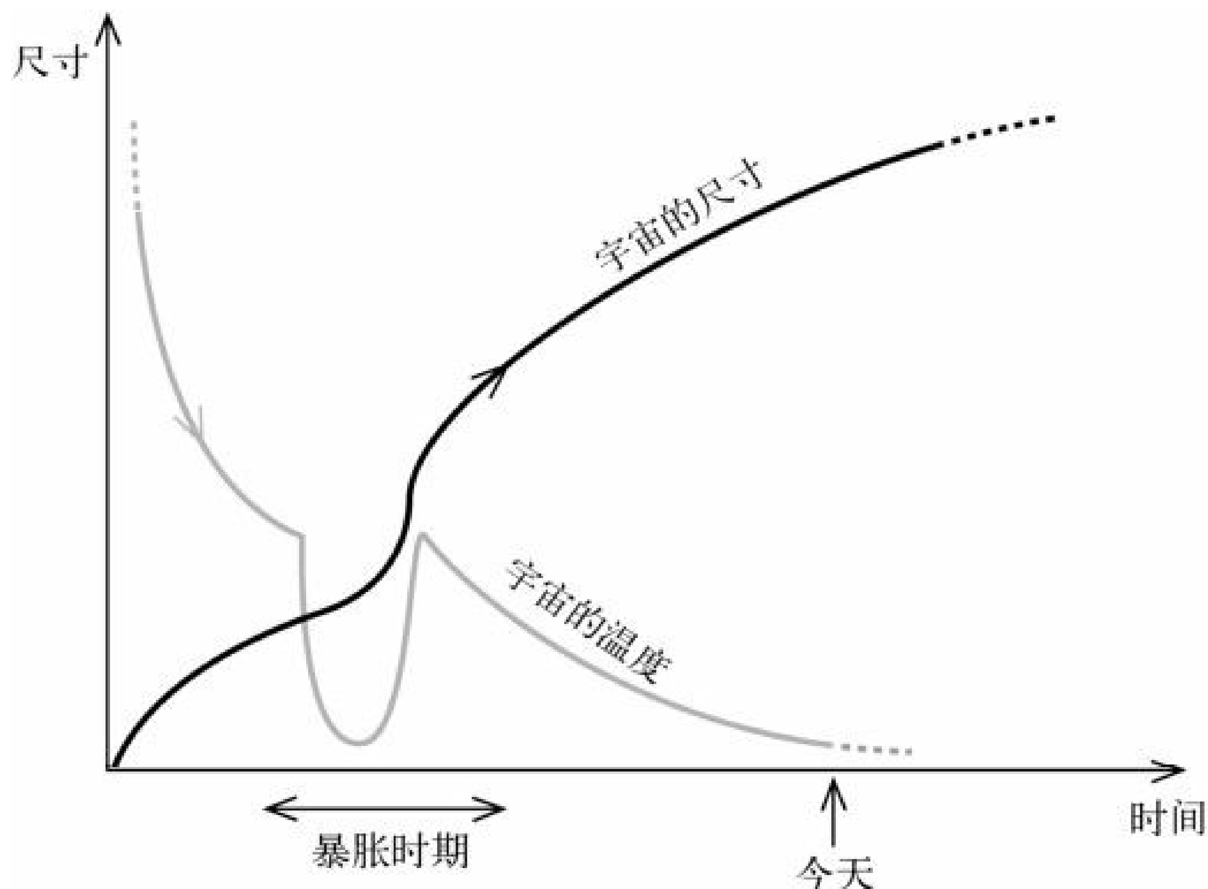


图 9.2 暴胀使得宇宙的膨胀突然加速，同时宇宙的温度急剧下降。暴胀停止时，驱动暴胀的粒子会发生衰变，在一瞬间产生了大量的能量。这导致宇宙被重新加热。然后宇宙又开始冷却，而且冷却的速度跟暴胀之前冷却的速度一样

粒子物理学的新进展为这个问题提供了一些线索。用粒子物理学家的语言来说，宇宙学常数只不过体现了宇宙的“真空能”——宇宙能够携带的最低能量。真空能也可以理解为一种局部上的最低能量状态，就像一个碗的碗底，如果你把玻璃珠扔进去，它就会滚到那个位置上。大量基本粒子都分别占据着这样的临时位置，每个位置都对应一定的能量。在宇宙冷却的过程中，粒子间的相互作用使得这些粒子都会随机变动各自的位置（图9.3）。

如果宇宙中的某些物质偏离了平衡状态，并开始朝着一个新的、能量更低的平衡状态移动，那么这种移动就会引发真空能产生排斥性的引力，从而导致宇宙的加速膨胀。[17]古斯意识到，如果这个过程发生得太快，就会产生灾难性的影响。宇宙的各处会同时产生大量新的真空泡沫。泡沫会迅速膨胀，相互碰撞。结果导致宇宙各处都产生了大量不规则性，满地狼藉，密度和膨胀速率变得极为不均匀。这不是我们所生活的宇宙。

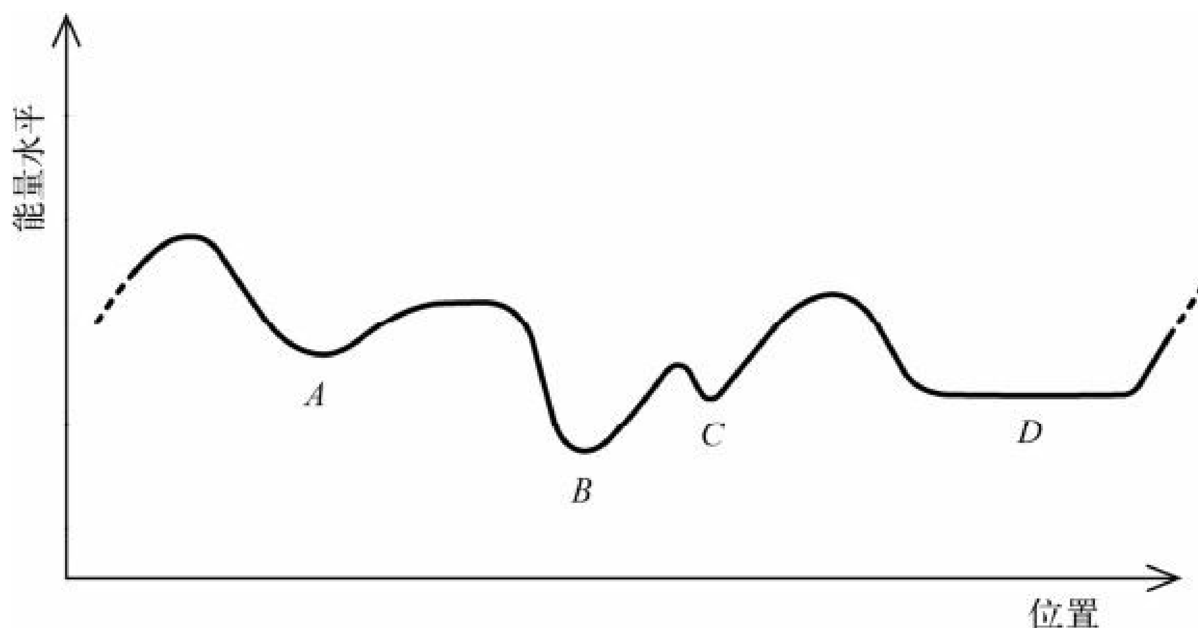


图9.3 早期宇宙包含的物质可以存在许多不同的能量极小状态，或者叫“真空”，每种真空的能量都不一样。在这张图中，A、B、C、D四个点都各自表示一种真空。真空态 B包含的能量最低，所以如果宇宙的一

部分处于真空态A或者真空态C的话，一个扰动就有可能使其翻过“山岗”，落入真空态B之中

尽管如此，古斯点出了这个问题之后，还是发表了他的暴胀理论，因为这个理论有一大堆非常简单而又令人满意的结果。很快就有人发现，宇宙从一个真空态向另一个真空态的转化可以进行得很缓慢，这样就能保证我们的可见宇宙处于其中一个新的真空泡沫之内。这就意味着在我们的可见宇宙中，没有发生过泡沫的碰撞，也没有碰撞的残骸。如果宇宙甚早期存在能产生排斥性引力的物质，而且所产生的排斥性引力满足正确的形式的话，这种排斥力就会驱动宇宙加速膨胀。但是它又会立刻衰变成普通的辐射和物质，于是宇宙就停止了加速，又回到减速的过程中来。

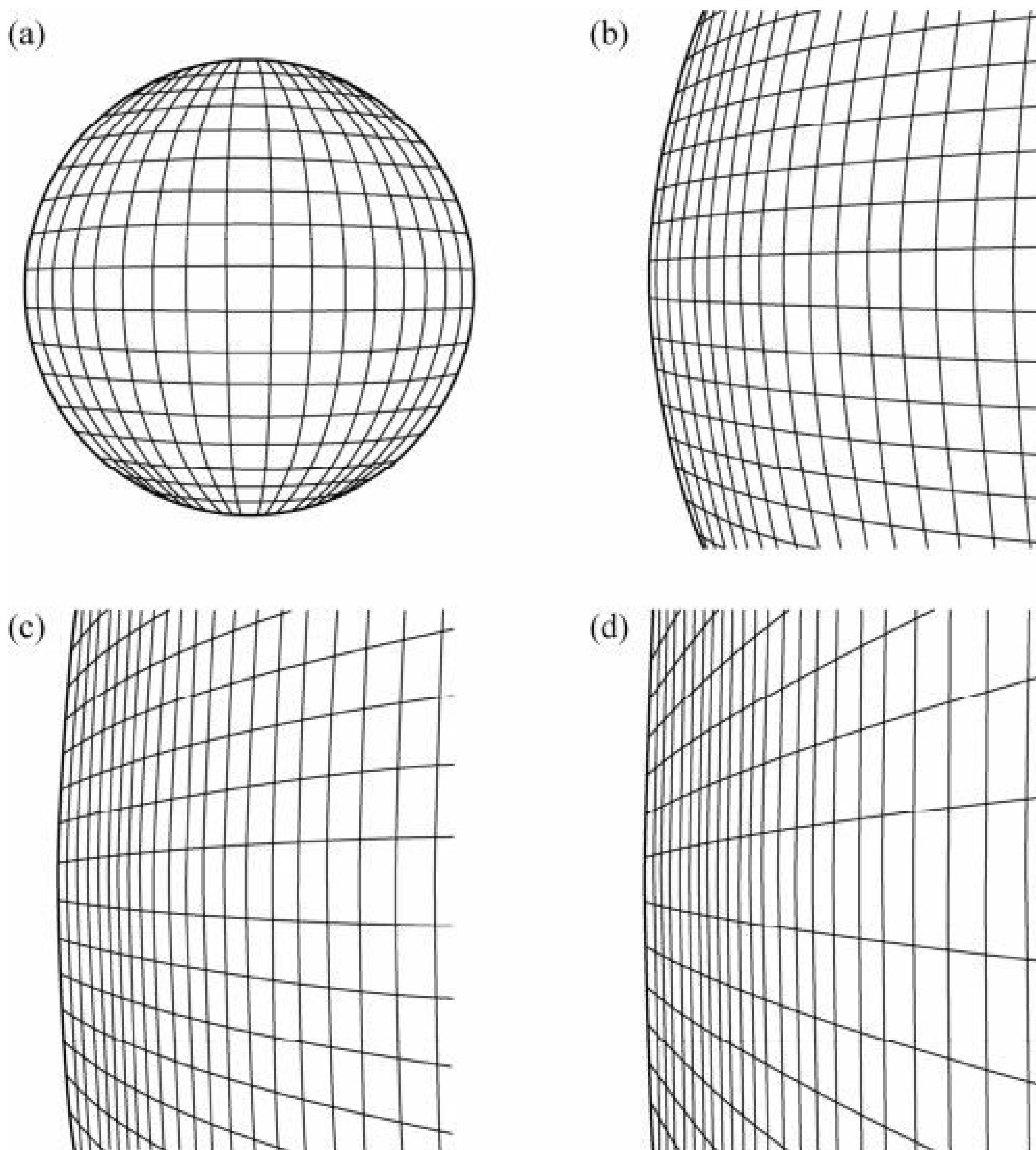


图9.4 一个曲面所经历的暴胀。可以看出从(a)到(d)，暴胀使得这个曲面从局部看越来越平坦

如果宇宙经历过这样的插曲，貌似就能解决宇宙学中存在已久的问题。此时，宇宙的体积将会更大，膨胀的速度也将会更快。在这个过程中，宇宙的膨胀处于一种临界状态，介于永远膨胀的宇宙和最终要回到大塌缩的宇宙之间。与此同时，宇宙变得非常均匀，空间的不同方向趋于一致（图 9.4）。先前的理论并不能解释可见宇宙的这些性质。短暂

的暴胀马上能将这些问题一股脑儿讲清楚。

暴胀理论最有意思的结论是，比以前设想的远少得多的物质和能量的原初涨落，通过加速膨胀，就能形成今天宇宙的可见部分（直径超过140亿光年）。这个涨落必须很小，不影响时空的光滑性，以使光线能够顺利穿过。[18]我们从宇宙中所观察到的高度的均匀性，正是由于可见宇宙源自暴胀之前的微小涨落，在膨胀的过程中，光子不断将能量从相对高温的区域带到低温区域，这就保证了空间的均匀和各向同性（图9.5）。

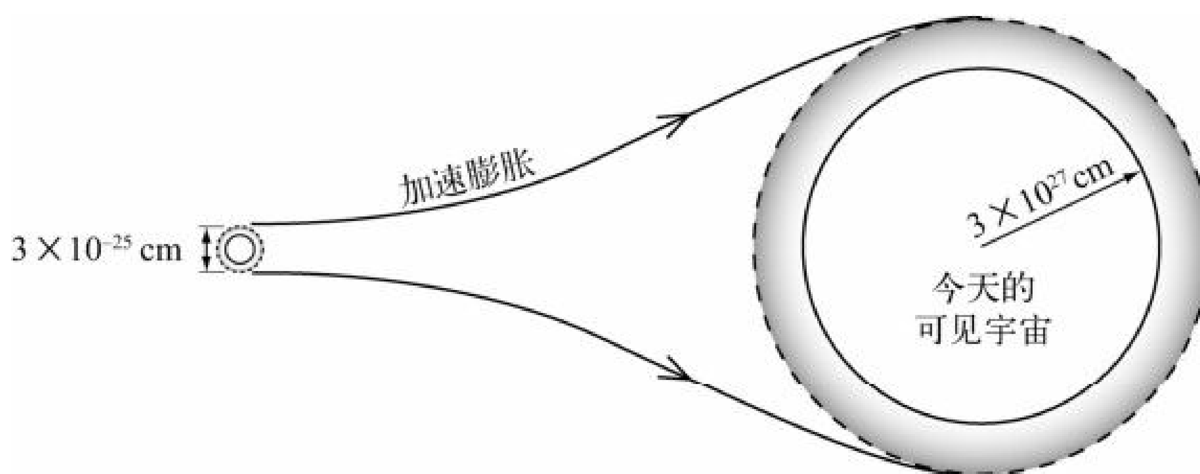


图9.5 暴胀使得一小片空间变得比今天的可见宇宙还要大（3×10^{27}厘米）。[①]这片空间必须非常小，这样在早期宇宙的任何时候，光信号都来得及将其中的不均匀性抹平。在这张图中，我们将早期设定为10^{-35}秒。所以光信号以3×10^{10}厘米每秒的速度传播，在这段时间内，就只能传播 3×10^{-25}厘米，这就是这一小片空间的大小。当时，宇宙的温度是3×10^{28}开，而今天已经下降到了3开。温度下降了10^{28}倍，同时暴胀将那一小片空间从3×10^{-25}厘米膨胀成了 3×10^{27}厘米。这就解释了为什么可见宇宙大体上是均匀的：我们看到的是一小片均匀的空间膨胀后的样子。如果宇宙没有经历过暴胀，那么可见宇宙就得从一片3厘米长的空间开始膨胀，而这比宇宙早期时光信号所能传播的最远距离还要大10^{25}倍

① “卡斯纳博士的宇宙”一节说可见宇宙的范围是10^{29}厘米。考虑到宇宙的膨胀，光在 137亿年间传播的范围大约有930亿光年，约为8.8×10^{28}厘米，所以10^{29}厘米一说更合理。——译者注

最后一点也很重要，暴胀同时也解释了磁单极子的问题。之所以会产生磁单极子，是由于不同方向上的磁力指向不匹配。由于可见宇宙原先对应的区域非常小，刚好没有出现磁力不匹配的现象，因此我们也就

不指望宇宙中存在磁单极子了。相反，如果宇宙先前没有加速膨胀过，宇宙原先对应的区域就会比暴胀宇宙模型中的大了约10^{25}倍。这样的话，宇宙中就会充斥大量磁场不匹配的现象，于是磁单极子就会泛滥成灾，光的平滑能力就被局限在了约10^{-25}厘米的范围之内。

暴胀解释了一揽子问题，包括可见宇宙为什么是均匀而各向同性的，为什么没有充斥大量磁单极子，为什么宇宙的膨胀处于临界状态，既不能保证会永远膨胀下去，又不能确定将来会开始收缩，最终回到密度极高的状态。暴胀理论能够以各种有趣的方式同我们曾面临过的问题和研究过的宇宙模型联系起来。从它身上可以看出一些稳态宇宙的影子，只不过这种"稳态"是暂时的。这一点很重要。如果暴胀不结束，加速永远持续下去的话，那么所有的一切就都会被抹得无穷均匀，宇宙中也就不可能存在星系和恒星了。

暴胀理论面临着与混沌宇宙模型相同的挑战。暴胀理论成功地解释了宇宙的高度各向同性和高度均匀性，不像混沌宇宙那样要求不规则性消失，也不需要解释因此而产生的其他问题。暴胀将不规则性驱赶到了今天宇宙的视界之外，而不是将它们斩尽杀绝。在很远很远的地方，这些不规则性依然存在。不过，由于整个可见宇宙都是由一小片区域经过暴胀形成的，所以还是体现出了该区域高度的均匀性和各向同性。

早先，米斯纳和其他"混沌宇宙学家"的努力从未成功过。他们无法证明，宇宙膨胀开始时的各向异性能够迅速减弱，以符合大爆炸核合成过程预言的氦元素丰度，以及微波背景辐射温度谱的高度各向同性。如果开始时各向异性的程度过于严重，为了消除这些各向异性就会产生过多的热辐射。更糟糕的是，有些类型的不规则性（存在于弯曲空间中的那些）貌似对这种简单的削弱机制有很强的免疫力。宇宙学家们长期以来相信一个看上去非常合理的假设，即早期宇宙的引力始终是吸引的。他们并不排除宇宙的晚期会出现排斥性引力，就像爱因斯坦曾提出的那样，宇宙学常数可能是正的。在1980年的时候，否定这个假设虽然不太可能，但毕竟还是存在这种可能性。如果宇宙早期也存在这样的排斥性引力的话，所有钉子户般的不规则性都会被迅速抹平，因为膨胀的速度太快了。

混沌暴胀的宇宙

要是你们能够洞察时间所播的种子，知道哪一颗会长成，哪一颗不会长成……

——莎士比亚[19]

暴胀宇宙学和混沌宇宙学都认为，宇宙的初始状态非常有可能是混沌不堪、一团乱麻的（相较于简单的状态，混乱的方式千变万化，因而有更多的可能性出现），而其中如果有一小片区域开始暴胀，体积就会迅速增大，空间就会变得更光滑、更具有各向同性，最后会将其他暴胀区域都挤到了视界之外。有一点非常重要。暴胀宇宙学不能从整体上描

述宇宙的行为。我们不知道视界之外的那片宇宙空间是否都很平滑，或者是从头乱到尾。

就连这一小片区域也不会是绝对光滑的。其中必然存在着极为微弱的统计涨落和量子涨落，这些涨落最终会变成我们今天观察到的宇宙尺度的密度变化，例如星系。如果不是暴胀将它们放大，这些初始统计涨落的强度就太弱了，无法利用引力的不稳定性形成星系。宇宙学家们之所以会认识到这一点，是因为1982年6～7月的两个星期间，剑桥大学举办了一期关于这个新理论的特别研讨会。[20]

与会人员很快就认识到，暴胀产生的涨落遵循一种很特殊的模式。过去，人们一直在研究这种模式的涨落，因为它最简单最方便。当这种模式的涨落被放大后，它就会使宇宙所有尺度下的结团行为拥有大小相等的强度。[21]这确实很特殊。考虑了不规则性之后，还要让宇宙大体上永远保持弗里德曼–勒梅特宇宙的模样，满足这个要求的模式是唯一的。

我们之所以这么说，原因在于宇宙的暴胀时期可以暂时地看作是一种稳态宇宙。想一想在稳态宇宙理论中，德希特的指数暴胀使得你无法区分宇宙的过去和未来。如果某种机制导致了微弱的不规则性，那么你也必然无法利用这些不规则性来区分过去和未来。如果在一个假想的观测者看来，任何尺度的涨落都得有相同的强度，那么以上的原则就不会被打破了。

因此，作为意外之喜，暴胀还能解释星系的存在，解释这种特殊的密度不规则性。暴胀会导致宇宙微波背景辐射的温度谱上出现一种特定的与角度相关的变化。我们得做一个实验，才能知道暴胀是不是真的发生过。

宇宙学家们开始用一系列观测仪器仔细地寻找这个铁证，结果大大支持了暴胀理论。美国国家航空航天局（NASA）先后发射了两颗人造卫星，一颗是宇宙背景探测者（COBE），另一颗是维尔金森各向异性探测器（WMAP），用以探测宇宙背景辐射是否存在这种特殊的印记，它的温度是否会随着方向的不同而发生变化。相比地面上的仪器，人造卫星的探测效果更好，因为它们不需要穿透变幻莫测的地球大气，而且能够扫描整个天空，获得海量的温度差分数据，这样的话，就能把任何数据采样中都存在的偶然性变得微乎其微。

宇宙学家们所要探测的是，在相差一定角度的两个方向上测量到的温度差异有多大。标准的暴胀宇宙模型所预言的结果由图 9.6 中的实线表示，其中相差的角度涵盖了很大的范围。此图的特征十分明显。随着相差的角度越来越小，振荡也在衰减，就像袅袅的钟声。越向右延伸，曲线所对应的空间尺度就越来越小，最终一切涨落都被熨平了，这是由

于能量不断地从高密度区域流向低密度区域。如果将这条实线向左延伸，起伏就没那么频繁了。这就与老式的COBE卫星的观测结果符合得非常好，这个卫星只能比较相差十度以上的两个天空区域的温度差异。与实线精确重合的数据点主要来自于WMAP卫星的结果。[22]我们注意到，在第一个拐弯处，观测数据与暴胀理论的预言分毫不差，但是当我们趋近仪器灵敏度的极限时，观测结果的不确定度就增大了。数据信号向角度很大的区域行进时，会遭遇一次奇特的“下沉”，对此天文学家已经展开了大量讨论，并提出了各式各样的解释。[23]

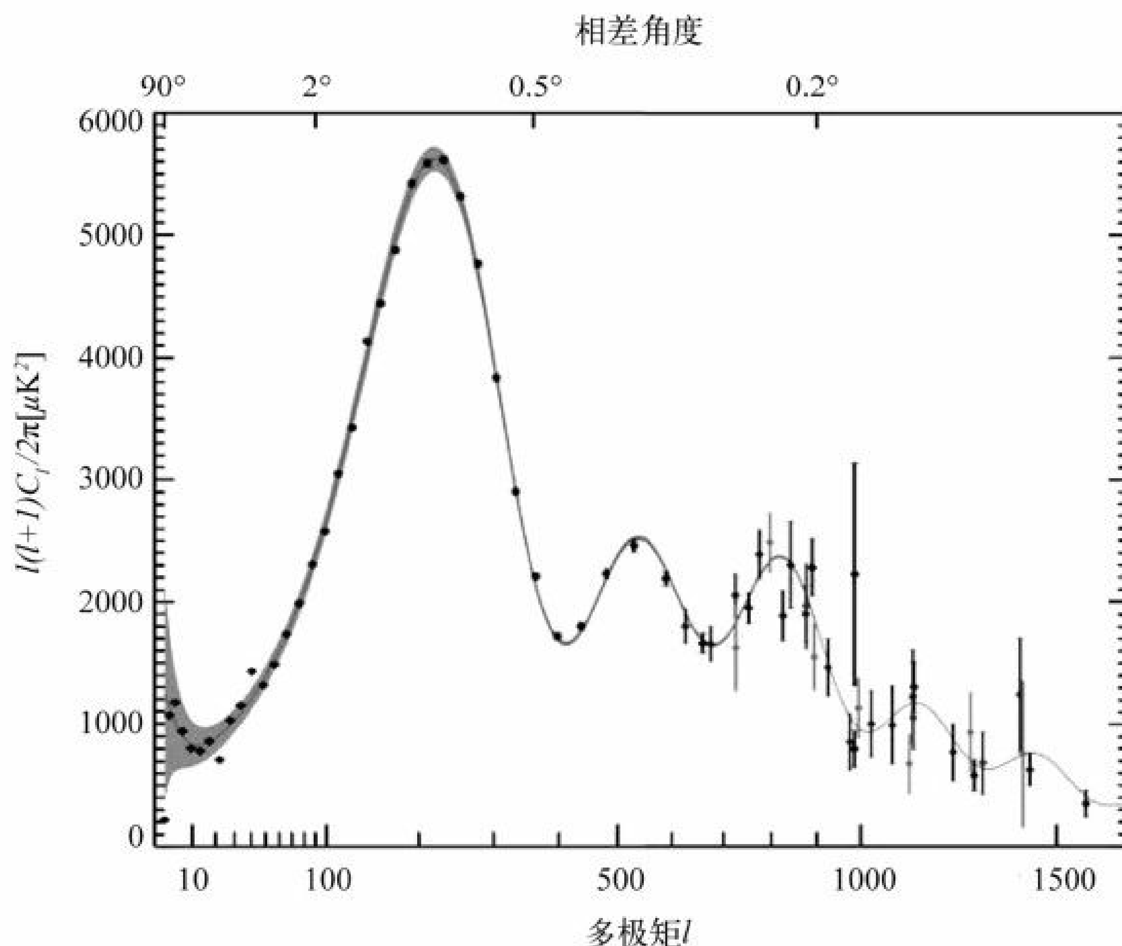


图9.6 暴胀宇宙理论预言的、微波背景辐射的温度涨落水平（以微开尔文的平方为单位）随相差角度的不同而发生的变化（实线），观测数据来自于地面实验或气球实验。作为比较，满月时月亮的大小是相差0.5度。相差角度很大的数据主要来自于WMAP卫星服役七年中所采集的数据（黑点）。左边数据上的深色带表示无法消除的统计不确定性，因为可见宇宙中这样大尺度的区域数量有限

2009年夏，一颗名为“普朗克”的新卫星被欧洲空间局发射升空。不久，它就会传回更加精细的观测结果。同时，地面的天文学家一直在关注电子科技的更新换代，希望能在这轮竞赛中造出更灵敏的探测器，获得更完整更精确的辐射信号，以便更深刻地了解宇宙历史的第一个瞬间。暴胀真的发生过吗？也许图 9.6 最终的样子会说服宇宙学家相信这一点。它揭示了暴胀的旋律，让我们能够回溯宇宙诞生10^{-35}秒后的一刹那。作为现代宇宙学最新进展的代表，它常常出现在科学新闻、科学讨论和公众讲座当中。总有一天，它会被看作描述人类所能了解的宇宙最早时刻的第一份证据。它就像一张婴儿宇宙的照片。

永恒暴胀的宇宙

世界不足吾欲

——詹姆斯·邦德家族的座右铭

人们完全没有想到，描述早期宇宙的暴胀理论还产生了两个不同寻常的推论。暴胀理论提出，早期的宇宙经历过一次短暂的加速膨胀。于是，我们今天所能看到的那部分宇宙空间，就是由宇宙早期某个不起眼的角落膨胀而成的，这个角落必须足够小，以使光线和其他物理机制来得及将它熨平。起初，我们几乎完全被这个简单的理论迷住了，它成功地解释了我们所看到的宇宙的总体性质：大体上均匀，其中洒满了不规则性的种子，这些种子注定会变成星系；它的膨胀速率恰到好处，还有高度的各向同性。从前，这一切都是无法解释的巧合。而如今，从一个简单的假设就能推得所有结果。

然而，这个简单的图景很快就凸现出一些意想不到的麻烦。说原初宇宙的一小片空间均匀地膨胀成了我们所看到的这部分宇宙，完全没有问题——但是旁边的空间后来怎样了呢？每一个这样的空间都可能经历一次大同小异的暴胀，各自变得巨大而均匀，但最终和我们这部分宇宙的性质并不相同。我们看不到这些空间，是因为光线没有充足的时间飞到我们这儿。[24]但在万亿年后的某一天，我们的后代或许会发现，世上还存在这样一个空间，它的地理环境与我们的宇宙截然不同。我们估计，从超大尺度上看，整个宇宙异常复杂，毫无规律，虽然从我们今天所能看到的尺度上来看，它又显得极为均匀、相对简单。暴胀理论告诉我们，我们所能看到的局部宇宙并不能体现整个（也可能是无穷大的）宇宙的性质。宇宙的地理环境十分复杂，比我们所能想到的还要复杂得多（图9.7）。

如果这还不够可怕，那还有更可怕的事情。艾力克斯·维连金（Alex Vilenkin）和安德烈·林德（Andrei Linde）发现了暴胀宇宙的另一件麻烦事：它会自我繁殖（图 9.8）。宇宙的一部分空间一旦开始暴胀，就会加速膨胀，并且会导致其中一部分子空间进一步暴胀。这是一个自我繁殖的过程，一次暴胀诱发了更多暴胀，子又有子，子又有孙，

子子孙孙无穷匮也。这就叫“无限”。如果未来是无限的，那么过去怎么可能不无限呢？

自我繁殖的永恒暴胀意味着，尽管我们的小宇宙就像一个暴胀的“泡泡”，可能存在一个开端，但整个“多重宇宙”（multiverse）却不存在任何开端，也不会有任何终点。我们生活在其中一个（或许很罕见的？）泡泡之中，这个泡泡膨胀了足够长的时间，使得恒星和行星能够产生，生命得以演化。宇宙的历史也十分复杂，比我们所能想到的还要复杂得多。

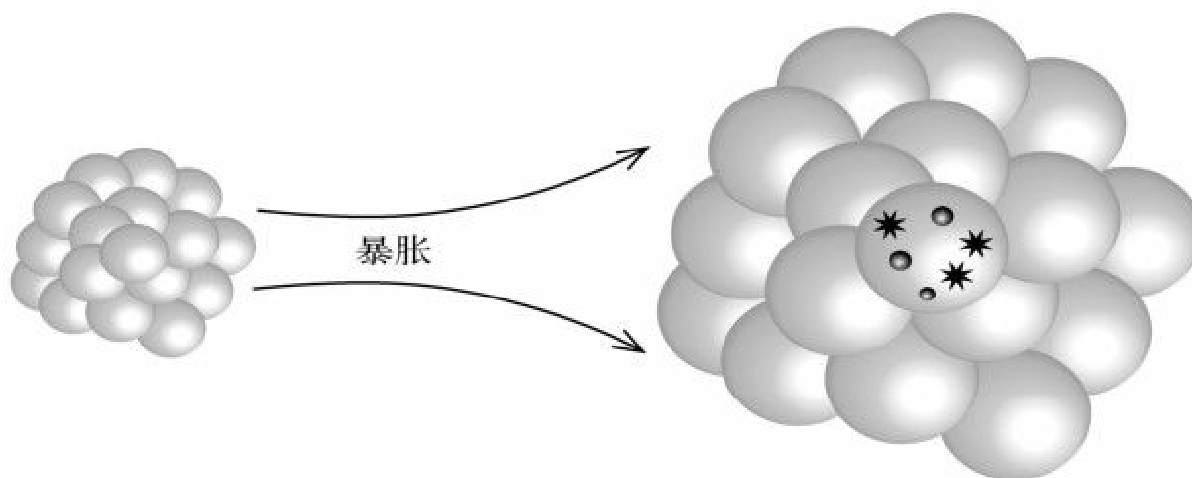


图9.7 混沌暴胀。早期宇宙的部分角落随机地经历了不同程度的暴胀。我们发现自己生活在其中一个子宇宙中，体积足够大，寿命也足够长，这样才能形成恒星，才能演化出碳基生命[25]

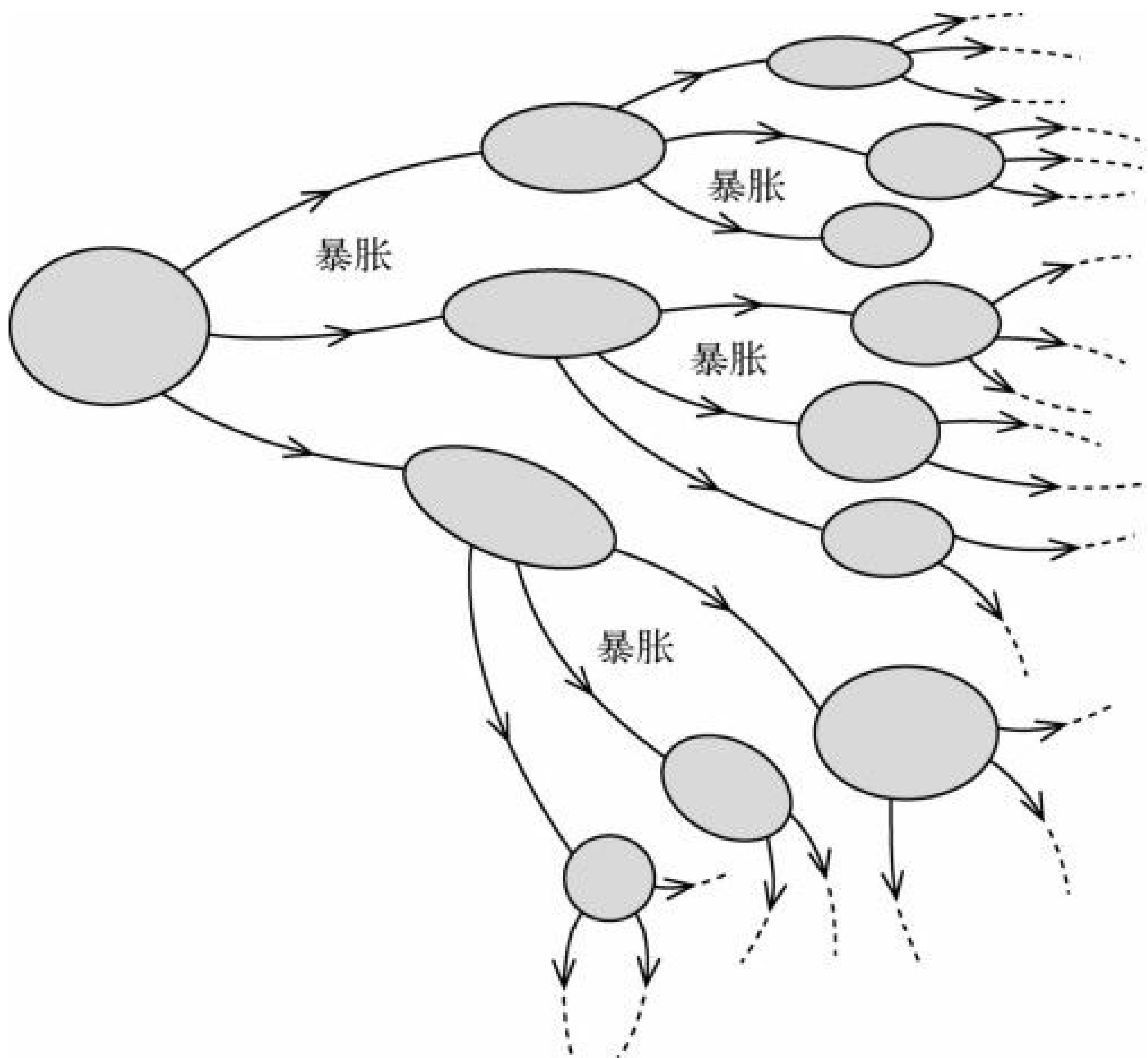


图9.8 自我繁殖的永恒暴胀。当一小片空间发生了暴胀以后，就会为其中某些部分创造了将来进一步暴胀的条件。这个过程看起来没有尽头，也许也没有起点

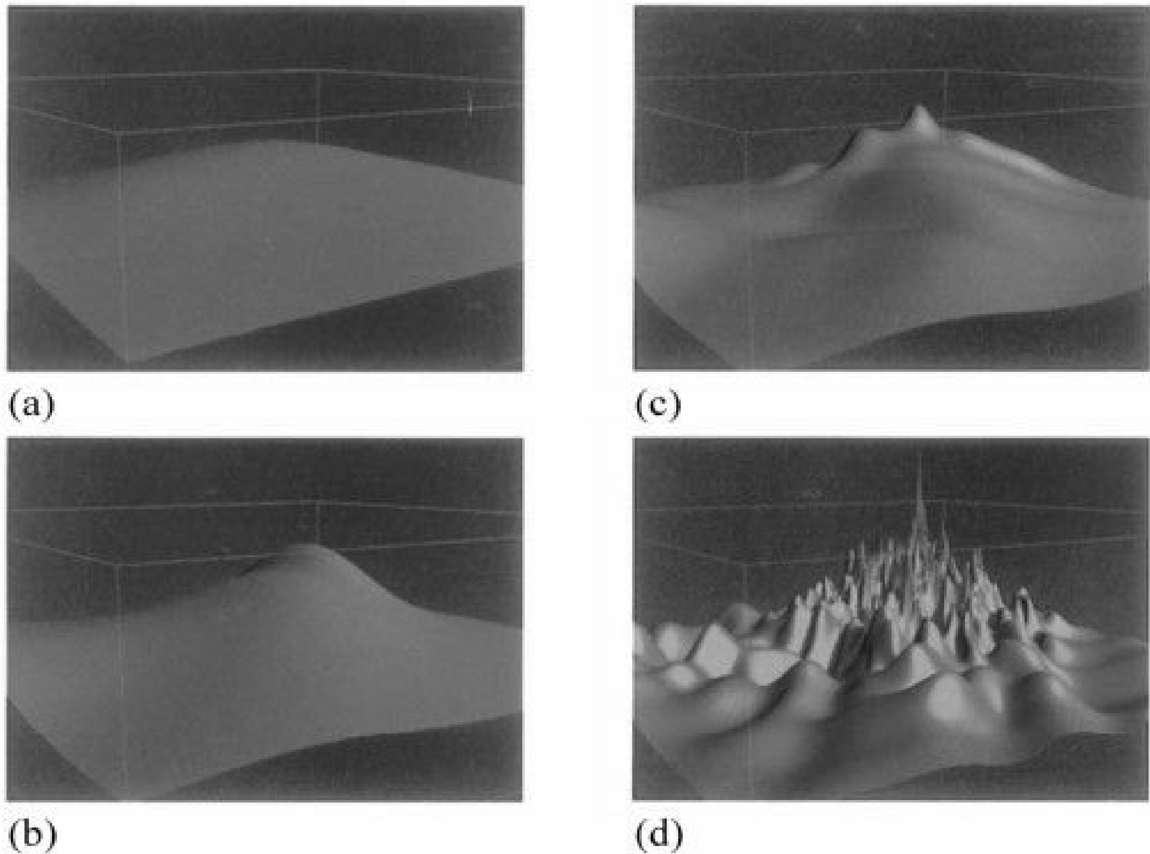


图9.9 从(a)到(d)分别是安德烈·林德和迪米特里·林德用计算机模拟的不同阶段的自我繁殖暴胀的截图。空间中的小山头表示暴胀的开端。后来像石笋一样的山上有山，表示自我繁殖的暴胀宇宙已经脱缰狂奔了

这种“永恒暴胀”的理论使我们对宇宙历史的认识变得愈发复杂。这一点十分恼人，因为当我们意识到暴胀宇宙的地理环境异常复杂时，又会发现我们所生活的这个宇宙的历史也是多姿多彩、错综复杂的，其中的大部分内容永远无法为我们所知。宇宙就像一个精致的被子，而我们只是生活在其中一块简单而独特的时空补丁上。

永恒暴胀理论的创始人之一安德烈·林德，生动地描绘了自我繁殖的发生过程（图 9.9）。图中的起伏逐渐变成小山，表示宇宙中的不同区域在以不同的速度暴胀。这张计算机模拟图由他的儿子迪米特里绘制。而那些山头上的尖峰以及尖峰上的尖峰，表示的是随着暴胀不断地进行自我繁殖，空间的复杂性也与日俱增，形成了一种分形结构。我们生活在其中一个独特的尖塔上[①]，暴胀早已偃旗息鼓，膨胀也已趋于平静。但是我们所处的这个位置并不典型。无限重宇宙的大部分空间仍处于暴胀之中。暴胀形成的这幅丰富多彩的石笋图，被林德称为“康定斯基的宇宙”［不过，这更像是美国概念艺术家所罗门·莱威特（Sol

LeWitt）的雕塑《污渍，第 15号》（Splotch 15），如图 9.10］。这是一张表现宇宙分形永不停止的快照。

①应该说我们生活在其中一个低谷中，因为在这张图中，高处表示暴胀仍在继续的区域，低谷表示暴胀已经结束的区域。——译者注



图9.10 所罗门·莱威特在2005年的雕塑作品《污渍，第15号》。这个概

念艺术装置高3.66米，用丙烯酸树脂玻璃纤维制成。这是他自2000年开始创作的相似风格的系列结构中的一个（共22个）

突然间，我们的宇宙又变简单了许多

电子报表造就了一个“假如这样假如那样”（what-if）的社会。这种社会里的人不会像常人一样立即付诸实践，而是对每一步行动都满腹狐疑。他们对任何事都会再三斟酌。

——约翰·C.德沃夏克（John C. Dvorak）[26]

暴胀宇宙理论看似简单却大获成功。它解释了我们所看到的宇宙的本性，同时，它又试图告诉我们，视界之外还存在看不到的无穷复杂的广阔空间。支持微波背景辐射中存在特定温度涨落的观测证据越来越多，这意味着我们要认真地对待暴胀理论，我们所看到的这部分宇宙在很早的时期经历过一轮暴胀。这解释了为什么我们会发现宇宙如此均匀且各向同性，为什么其膨胀速度会接近临界值的大小，为什么它不包含磁单极子，以及为什么它又有特定的不均匀性散布其中——在牛顿首先提出的、栗弗席兹证明了的、爱因斯坦理论中也存在的引力不稳定性的作用下，这些不均匀性最终会长成一个个星系。

暴胀理论跟勒梅特八十年前所研究过的一个宇宙模型很像。区别之处在于，暴胀理论构造了一个变化的宇宙学常数，它可以迅速衰变成辐射，并留下了可观测的痕迹。这简直妙不可言。宇宙的最初 10^{-35}秒时所发生的一些事变成了的化石，直到今天还能看到。为了了解在这个早得出奇的时期发生了什么事情，我们得以有机会将理论预言与天文观测进行比较。

但除了成功地解释我们所看到的宇宙，暴胀理论还是吓了我们一跳，因为它得出了一个看似必然的结论，即我们只是生活在无穷复杂的一堆膨胀“泡泡”的一个当中，每个“泡泡”的结构可能千差万别，统治其中的自然法则也可能各不相同。突然间，不止只有一个宇宙，而是有了很多很多。

宇宙也疯狂

据我的发现，世界上没有灾难，只有机会，事实上是发生新灾难的机会。

——鲍里斯·约翰逊（Boris Johnson，现任伦敦市长）[27]

暴胀的宇宙将空间中的微小涨落迅速扩充为整个宇宙，这使我们不得不把宇宙改写为复数形式。我们能够看见其中一个原初的微小涨落膨胀后的模样。然而，我们周围的一切，包括视界之外的空间，都是一轮又一轮暴胀自我繁殖的产物。我们所看到的这部分宇宙的模样取决于两个因素：一是这轮暴胀的具体细节，二是统治其中的物理定律。直到不久之前，大多数物理学家都还相信，世上只存在唯一一种“万有理论”。寻找这个理论的过程就像在玩一种大型拼图游戏。每个地方都只有唯一一块合适的拼图，但你不知道什么时候才能找到它。

但渐渐地，人们对于这个简单预期的信心丧失殆尽了。种种万有理

论的最佳候选理论具有完全不同的性质。这些理论中都存在一种整体的对称性，这种对称性限制了自然法则可能具有的形式，同时它们又具有一种程度出人意料的灵活性。可以存在各不相同但自洽的物理定律，其中的相互作用数目可以不同，允许的空间维度可以不同，还有很多其他性质都可以不同。许多从前人们认为宇宙固定不变的性质，现在看来却是随机产生的结果。我们从自然界中看到的相互作用力，总结出的定律法则，只不过是其中一套可能的地方性法规罢了。除此之外，其他地方的定律法则也都很完备，从各自的角度看都是自洽的。

起初，有能力角逐万有理论的候选理论似乎不太多，其中一类就是20世纪80年代早期出现的那几种“弦理论”。但到了后来，人们发现这些弦理论并不是终极理论，它们并不是真正的万有理论。实际上，这几种弦理论是一种更深层理论的几种极限情况，这种未知的万有理论叫做“M-理论”。我们只知道世上应该存在这样一个理论，但对细节一无所知，只知道它存在几种极限情况，例如能量和温度很低的时候，或者引力很弱的时候。值得注意的是，这种理论预言了一种相互作用，很像爱因斯坦的广义相对论所描述的引力。据我们所知，这种理论极为复杂，包含了海量的自洽世界；乐观估计，一共有超过10^{500}个世界。这个数目实在太大了。

在10^{500}种可能的状态中，都存在各自的定律和物理常数。当我们这一片宇宙冷却到可以开始暴胀时，所处的状态只不过是其中之一。这个充满了可能性的国度被称作“弦景观”。[28]在这些可能存在的世界中，并不是只有密度、温度或者膨胀速率之间存在微小的差异，就像我们在这本书中所提到的那些宇宙模型一样。它们从根本上就是千差万别的：相互作用的数目不同，自然常数的数目不同，就连空间和时间的维度也不同。其中许多世界都可能缺乏像电磁力之类的相互作用，于是就不可能存在原子或生命。

对这些海量可能性的研究仍然很原始。后来，一种叫做卡拉比–丘流形的复杂数学结构被人们引入，促进了相关研究的发展。数学家尤金尼奥·卡拉比和丘成桐很早以前就发现了这种流形，比物理学家的研究早得多。[29]起初，这些研究看起来很吓人。我们对卡拉比–丘空间的了解不多，图9.11是其中一幅图片。这些空间可以由一些表示大小和复杂度的特征量来分门别类。

研究人员充分利用了高速计算机去研究这些可能性，并将它们分类整理，但人们力所能及的也不过是弦景观的一些稀疏角落，而且其中的事物也不能太复杂。这些研究燃起了一丝希望，或许在卡拉比–丘流形所定义的巨额数目的空间中，只存在少量“有意义的”世界。其中一些世界中的物理规律可能会产生像电子、夸克之类的粒子，进而组合成原子

和分子。

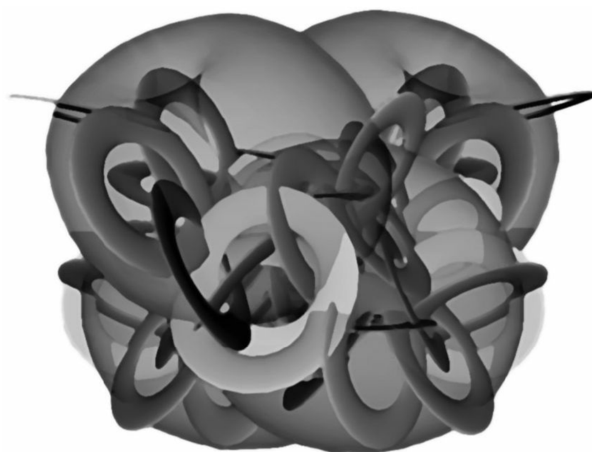


图9.11 一种卡拉比–丘空间。

同时，我们还面临一个巨大的宇宙彩票问题，永恒暴胀宇宙的不同角落会随机地落入10^{500}种弦景观状态中的一种。该角落的暴胀结束后，随机产生的状态将会决定暴胀产生的巨大空间的性质。

弦景观所包含的物理规律的数目听起来相当惊人，即使只研究其中一部分卡拉比–丘空间，难度也相当大。这让数学家们忙活了好一阵子。然而，罗格斯大学的弗里德里克·德内夫和迈克尔·道格拉斯指出，这种研究弦景观的系统性方法存在一个严重的问题。[30]有的人可能会想，物理学家们所采用的策略是，寻找弦景观中能演化为我们所观察到的宇宙的地方。可惜，这个计划过于理想化，没有考虑到所谓的计算复杂性问题。仅仅计算弦景观中的最低能量状态就已经需要漫长的计算（对此的技术术语叫“非确定性多项式难题”，简称“NP 难题”），而研究所有这些可能状态的性质更是远远超出了任何人造计算机的能力范围（就连量子计算机也不行）。

这类所谓的“难题”，就是说随着信息输入量的增加，所需要的运算量就会呈指数增长。我们习惯于解决“简单”问题，其运算量与信息的输入量呈正比（或者是后者的某种幂函数）。计算机花了四十万小时（利用许多计算机同时计算），只是简单地扫描了其中一百万种模型。[31]但是复杂性和运算时间会随着“难题”范围的扩大而迅速增长，如果我们把模型的数量翻一番，运算时间就会长达一百多万年。

暴胀宇宙正在泛滥成灾！我们没法通过观测实验将不适合的候选模型逐个排除。这听起来十分让人焦虑，不过现实更加糟糕。到目前为止，我们只数出了弦理论存在多少种不同的物理规律，但还没开始计算永恒暴胀的宇宙会自我繁殖出多少个暴胀“宇宙”呢！

假设我们所生活的这一小片空间，随机选择了超弦景观的一个真空态。这就对应着一类特定的物理定律。这片空间开始暴胀时，体积最大不超过光速和宇宙年龄的乘积。为了满足我们在宇宙中看到的各向同性、均匀性以及磁单极子的稀缺性，暴胀必须进行得足够彻底，至少要让体积增大一个 $N=e^{60}$的因子。暴胀产生的副产品，也就是其他可能存在的“宇宙”至少有

$$e^{e^{3N}} = 10^{10^{77}}$$

个之多，这些“宇宙”的几何性质都不相同。这个数目实在太大了，给这些可能的世界列一张名单，产生的数据量都会多到无法被我们的大脑容纳。为了将10后面的10^{77}个零完整地写出来，就要把我们所看到的宇宙范围内的所有原子都变成墨水。这个数字比弦景观所包含的真空态的数目还要大得多。实际上，如果我们不单单考虑我们周围的这片宇宙空间，而是考虑更大范围内的宇宙所能产生的空间的话，这个数目甚至还会变得更大。整个宇宙永远在生成新的空间，这些空间数目惊人，一边膨胀，一边又在产生更多的空间，子子孙孙无穷匮也。

我们并不需要造出更大的数字，将这些信息完整地表达出来。可能的宇宙和真实的宇宙都多得难以置信，就像所谓的“可见宇宙”一样无边无际。

注释

[1] 在乔治和格拉肖最简单的理论中，质子最有可能衰变成一个正电子和一个中性的π介子，后者又会迅速衰变成两个高能光子。正电子的能量大约是460兆电子伏，两个光子则分别携带260兆电子伏的能量。正电子出射的方向和两个光子相反，而两个光子运动方向的夹角大约是40度。

[2] M. R. Krishnaswamy et al.,*Phys. Lett. B* 106, 339 (1981).

[3] K. Hagiwara et al.,*Phys. Rev.* D 66, 010001 (2002).

[4] 这些实验有时会遭遇挫折。1996年，日本建造了一个包含45 000吨水的超级神冈探测器，开始进行地下实验。这个探测器主要是用来探测质子衰变和中微子其他的新型相互作用。结果这个实验室对中微子的探测很成功，因此分享了2002年的诺贝尔物理学奖。不过，实验人员从来没有探测到质子衰变的迹象。2001年11月12日，灾难降临了。一个事故导致6600个用于记录衰变信号的光电倍增管发生了爆炸。多米诺骨牌效应随之发生，每一个光电倍增管的爆炸都产生了冲击波，击毁了周围的倍增管。每更换一个光电倍增管都要花2000英镑。更换光电倍增管的

工作一直持续到了2006年，不过现在已经好了。

[5] 这三个诱发正反物质不对称的因素是由著名的苏联武器物理学家、诺贝尔和平奖得主、人权活动家安德烈·萨哈罗夫（Andrei Sakharov，1921～1989）在 1966年的一篇意义深远的论文［*JETP Lett.* 5, 24 (1967)］中提出的。

[6] S. W. Weinberg,*Phys. Rev. Lett.* 42, 850 (1979); J. D. Barrow,*Mon. Not. R. Astron. Soc.* 192, 1P (1980); J. Fry, K. Olive and M. S. Turner,*Phys. Rev. D* 22, 2953 (1980).

[7] 宇宙中超过四分之三的原子是氢，氢的原子核只有一个质子。所以在某种程度上可以说，宇宙中的质子数量接近原子的数量。

[8] R. Browning, *'The Lost Leader, The Poetical Works of Robert Browning'*, ed. G. W. Cooke, Houghton Mifflin, New York (1899), p. 405.

[9] Y. B. Zeldovich and Yu. Khlopov,*Phys. Lett.* B 79, 239 (1978); J. Preskill,*Phys. Rev. Lett.* 43, 1365 (1979).

[10]值得注意的是，狄拉克在他一篇优秀的研究成果中提出，只要宇宙中存在一个磁单极子，就能解释为什么所有电荷都是某个基本电荷的倍数。参见：P. Dirac,*Proc. Roy. Soc.* A133, 60 (1931).

[11]实验物理学史上一桩咄咄怪事是，斯坦福大学物理系据称在1982年2月14日，也就是情人节那天，探测到了磁单极子。没有人能解释这个实验，也没有人能重复这个实验。人们高度怀疑这是实验人员布拉斯·卡布莱拉（Blas Cabrera）的恶作剧，因为如果实验结果真实存在的话，就说明宇宙中含有大量的磁单极子，而这与许多已知的实验和观测发生了矛盾。

[12]目前关于这个问题的综述，可参见：J. D. Barrow, 'Cosmology and Elementary Particles',*Fundamentals of Cosmic Physics* 8, 83 (1983).

[13] *Troilus and Cressida* I. iii. 345.

[14]因为他们考虑的是宇宙的甚早期，空间曲率项和宇宙学常数项在当时都可以忽略不计，看作是零。

[15]古斯将自己的故事写进了书里：A. Guth, *The Inflationary Universe*, Addison-Wesley, Reading, Mass. (1997).

[16]所谓标量，就像质量或温度那样，只有大小，它们的大小可以随着时间变化。所谓矢量，就像速度那样，既有大小又有方向。

[17]真空本身非常类似于爱因斯坦的宇宙学常数，因为这是唯一的一种无论观测者如何运动，对于所有的观测者来说都一样的作用。对于局部能量最低的真空态来说，这个条件必须满足。不然的话，如果相对运动能够导致它的能量进一步降低的话，它就不是真空了。

[18]如果一个宇宙学理论中没有暴胀，那么宇宙膨胀的进程就会减

缓，在宇宙早期小到可以来得及让光线穿过的角落，就不可能“发育”成现在的可见宇宙。如果不同的角落膨胀成现在可见宇宙的大小，那么宇宙各个地方的密度、温度和膨胀速率一定是千变万化，没有统一起来的。尤其是可见宇宙的直径大约是10^{27}厘米，如果不考虑暴胀，我们将它缩小10^{27}倍后，就只有 1厘米。这个缩小的比例对应于宇宙在最初的10^{-35}秒时的状态，当时的温度比现在高10^{27}倍。虽然当时的宇宙看起来很小，但是在这么短的时间里，光所能传播的距离只有3×10^{-25}厘米，比1厘米还要小得多。如果往理论中加入暴胀，宇宙就可以从早期3×10^{-25}厘米的状态膨胀成今天可见宇宙的样子。

[19]*Macbeth* I, iii, 58.

[20]这次研讨会的一份总结，参见：J. D. Barrow and M. S. Turner, ‘The Inflationary Universe –Birth, Death, and Transfiguration’,*Nature* 298, 801 (1982).这篇《自然》杂志的长文大概创造了某项纪录，因为这份手写的投稿在投出五天后就发表了。研讨会所取得的进展参见：G. Gibbons, S. W. Hawking and S. T. C. Siklos (eds.), *The Very Early Universe*, Cambridge University Press, Cambridge (1983).

[21]首先提出这一点的是爱德华·哈里森，参见：E. R. Harrison,*Phys. Rev.D* 1, 2726 (1969).后来它又多次被泽尔多维奇提到，参见：Ya. Zeldovich,*Mon. Not. R. Astron. Soc.* 160, 1P (1972).

[22] WMAP更多技术细节和最新论文，参见：http://map.gsfc.nasa.gov/m_mm/pub_papers/threeyear.html.

[23]这个现象可能并不具有统计意义，可能是因为与可见宇宙尺度相当的超大尺度涨落受到了某种限制。

[24]非常可能的情况是，我们的“泡泡”被膨胀了非常大的倍数，除了我们的可见宇宙，这个过程还熨平了我们看不到的更遥远的地方。如果不是这样的话，这将会个奇怪的、与哥白尼原理严重矛盾的巧合。这也意味着，我们要想接触到附近的一个条件迥然不同、甚至其中的物理定律也大相径庭的宇宙的话，就得等上很久很久的时间，甚至到时所有的恒星都早已熄灭了。

[25] J. D. Barrow, ‘Cosmology: A Matter of All or Nothing’, *Astronomy and Geophysics* 43, 4.9–4.15 (2002).

[26]转引自：John Naughton,*The Observer*, Business and Media Section, 18 March 2009.

[27]转引自：*Sunday Times* 4 May 2008, p. 15，这句话是他退出保守党影子内阁转任伦敦市长时说的。

[28]更确切地说，这些不同的状态都表示这种理论可能出现的“真空态”。如果宇宙选择了其中一种真空态而不是另一种的话，就会从真空

态那儿继承相应的性质和物理规律。如果真空态是唯一的话，那么自然界可能体现出来的基本性质也是唯一的。在弦理论中，真空态多得数不胜数。

[29] E. Calabi, *Proceedings of the International Congress Mathematician, Amsterdam*, vol. 2, Ervan P. Noordhoff/North Holland Publishing, Goningan/Amsterdam (1954), pp. 206–207; S. T. Yau,*Communications on Pure and Applied Mathematics* 31, 339 (1978); B. Greene,*The Fabric of the Universe*, Random House, New York (2004).

[30] F. Denef and M. Douglas,*Annals of Physics* 322, 1096–1142 (2007).

[31] F. Gmeiner, R. Blumenhagen, G. Honecker, D. Lust and T. Weigand, ‘One in a Billion: MSSM-like D-brane Statistics’, arXiv:hep-th/0510170.

第10章 后现代宇宙

如果在路上遇到岔路（fork），就把叉子（fork）拿走。

——约吉·贝拉（Yogi Berra，前职业棒球选手，多有巧言妙语）

随机的宇宙

［机械唯物主义者的］宇宙是人逃避现实、能把埋头进去的最小的洞。

——吉尔伯特·K.切斯特顿

自20世纪初以来，宇宙学家都在深入研究各种可能的宇宙模型，并与天文观测进行比对，然后从中挑选出一个与观测结果最为相符的。而现如今，他们要同时将所有可能性都考虑进来了——某种“多重宇宙”的可能性，其中每一个宇宙都占据多重宇宙的一个角落，每个角落都要比我们可见宇宙的全部还要宽广。这种想法的新颖之处在于，这些可能的宇宙都可以同时存在于现实世界之中，而不仅仅是哲学家口中的“可能的宇宙”、或然历史学家所谓的“或然的宇宙”，又或是奥运会银牌选手心中的“假如当初那样、结果就会那样的宇宙”。

我们回过头去看永恒暴胀的宇宙。在宇宙自我繁殖的过程中，新的宇宙不断破土而出，永无止境，而且这个繁殖过程似乎也没有开端。如果我们拥有上帝之眼，能够纵览全局的话，我们就会发现多重宇宙里的绝大多数宇宙仍然处于暴胀之中。个别角落就像我们所在之处一样，第一轮暴胀已经平息下来，又恢复了往日的减速膨胀。每个角落的形状、密度、温度和膨胀方式都各不相同。对于其中一些角落来说，暴胀的持续时间太短了，无法让宇宙的膨胀速度达到临界值；而对另一些角落来说，量子涨落被放大后产生了各式各样的不规则性，有的比我们所在之处的不规则性强得多，有的则弱得多。而且每个角落也都拥有各自的物理定律。

我们应该如何应对如此众多的可能性？可能存在的宇宙数目无穷无尽，用再强大的计算机也无法进行系统的研究。通过求解简单的方程，甚至是复杂的方程，我们都无法预言其中任何一轮暴胀的持续时间或真空能的大小：这些事都是随机的。

在不同的人看来，随机性的含义有所不同。有的人认为随机性就是缺乏秩序、杂乱无章。有的人认为随机性指的是不确定性，使人无法精确地确定某事。还有一些人认为，随机性意味着结局的变幻莫测，完全无法预测。在早期宇宙中，所有随机事件都源自于量子的不确定性，我们只能预测这种事件发生的概率。物质和能量的本质是量子的，有一种与生俱来的不确定性，这就是随机性的根源。收集的信息再多再详细也不能消除这种随机性。这种随机性就像我们描述世界时说的空间、时间

和运动一样，是一种内在属性。两个完全相同的起因并不能得出一样的量子结果。

甚早期宇宙中发生的随机事件波及了遥远的未来。密度的微小涨落最终形成了巨大的星系团，而它们都起源于宇宙暴胀时产生的量子不确定性和量子随机性。

这种量子的粗粒性意味着，如果我们想预测永恒暴胀宇宙中的事件进程，就必须学会概率论的语言。理想的情况下，如果你大体选定了一种宇宙（跟我们的可见宇宙非常类似），我们就很想知道，永恒暴胀宇宙在自我繁殖时，会有多大概率产生这样一种空间。可惜，到目前为止，宇宙学家仍然不知道如何回答这么难的问题，仍然不知道如何才能搞清楚其中所谓的“概率”或“可能性”的含义。[1]然而，这个问题还没有难到不可收拾的地步，人们提出了几种不同的解决方案，并分别进行了深入的研究。到目前为止，每一种方案都有不足之处。也许，我们离问题的解决只有一步之遥。[2]

概率性的宇宙

我一直在寻找用以表达现实的抽象方式以及会给自身启迪的抽象理念。

——埃里克·坎通纳（Eric Cantona，法国著名足球运动员）

在谈论可见的宇宙时，我们所提到的“不可能的”和“可能的”分别表示什么意思呢？假设我们已经得出了宇宙拥有某种属性的概率（例如其中存在原子或恒星的概率），形如图10.1所示。我们所画出的这张图很有代表性，描述了某个自然常数的概率分布。其中有一个峰，表示最可能出现的结果，从这个结果出发朝两边走，概率就会越来越小。这张图可能是多重宇宙理论的一个预言，于是我们想知道如何使用这张图来检验理论的正确性。我们的可见宇宙所拥有的性质在这张图中出现的概率是否为最大？更重要的是，如果我们的可见宇宙所拥有的性质在这张图中是极不可能出现的，那么我们是否该认为多重宇宙理论错了呢？

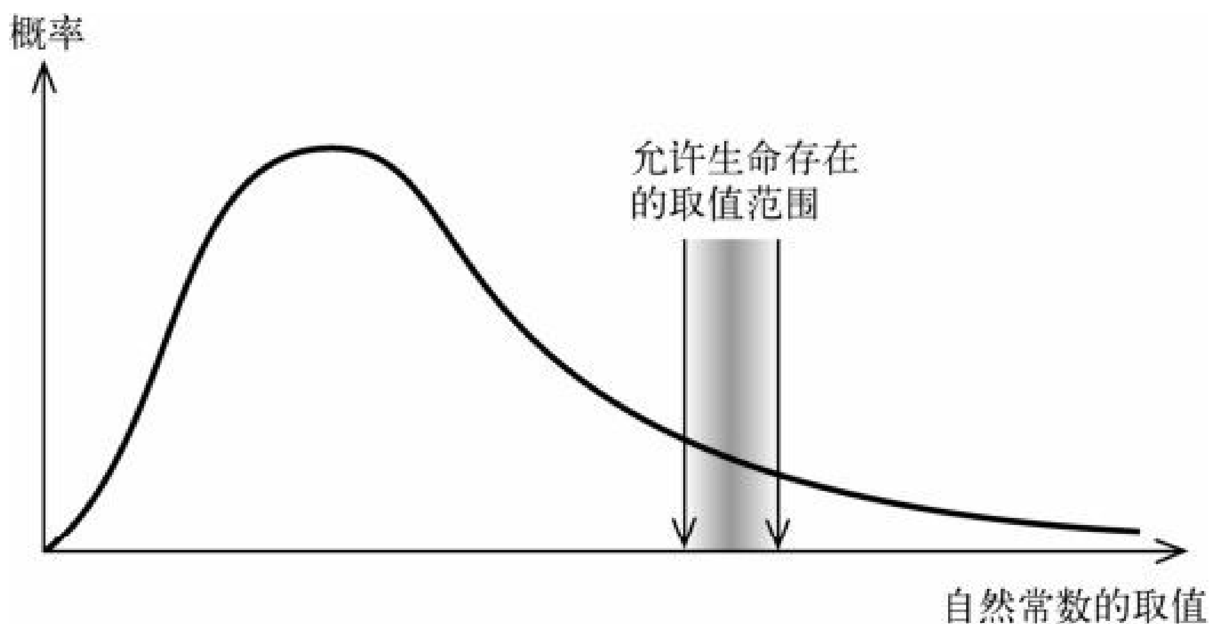


图10.1 宇宙中的自然常数可能拥有许多不同的取值，每种取值出现的概率各不相同。最可能出现的取值超出了允许生命存在的取值范围，因此我们观察不到这样的最可能出现的宇宙

对于以上问题，如果你回答“是”，那你就错了。为了明白其中的道理，我们需要了解一些多重宇宙理论中的“生命”知识。也许，任何形式生命都可以被狭隘地定义为“原子间的复杂行为”。要想使生命的出现成为可能，宇宙中的自然常数就必须落在特定的取值范围内。如果永恒暴胀最可能产生的结果处于这个范围之外，因而这种最可能的宇宙中不允许生命的存在（也许是因为该宇宙的温度从未降低到一百万度以下），那么我们就不该为自己连最可能出现的宇宙都没有见到而感到惊讶。此外，假如允许生命存在的自然常数的取值范围非常狭小，那么非常有可能的情况是，我们不能根据这张图就否定多重宇宙理论。重点不在于宇宙拥有特定属性的概率，而在于宇宙中可能存在观测者（“生命”）的同时，也拥有这种特定属性的条件概率[①]。至于不可能存在“观测者”（其中的观测者不必像我们人类一样）的宇宙，当我们用证据验证理论时，这种宇宙就不算数。

① 条件概率是在一件事已发生的前提下，另外一件事发生的概率。——译者注

一语惊醒梦中人。我们之前还不太习惯将宇宙学家的存在与否当作宇宙学理论正确与否的重要因素。

人择的宇宙

如果少一点自我的遮蔽，我们心灵的镜子便能多一点映照出宇宙。

——莫里斯·德鲁翁（1918～2009，法国小说家）[3]

宇宙学家早就意识到，宇宙的大尺度属性和生命是否能够存在之间存在某种惊人的紧密联系。乍一看，这件事有点儿怪。宇宙很大，数不尽的恒星和星系朝各个方向延伸了数百亿光年。这些东西怎么会跟此时此地、在一个不起眼的星系里绕一颗普通恒星转动的我们发生联系呢？

这种意外的联系之所以会存在，是因为宇宙的膨胀将空间和时间连在了一起。宇宙刚开始膨胀时，“生命”存在的基础，也就是能够形成复杂分子结构的化学元素，例如碳、氧和硅都不是现成的。宇宙最初三分钟的原初核合成仍未将那些元素制造出来，而只是产生了氘、氦和锂。不过，这些生命所需的元素都是经过百亿年的煅烧之后才被恒星制造出来的。这个过程涉及到一系列核反应，先是两个氦核合成了铍，再加入一个氦核就合成了碳，再加入一个氦核就合成了氧，依此类推。这种过程发生于恒星死亡的时刻，当死亡的恒星发生爆炸，形成一颗超新星时，所产生的化学元素就会散布到宇宙空间中去。最终，这些化学元素形成了宇宙尘埃、小石块，然后聚集成行星，进而组成分子和人类。

想要产生碳元素以及像太阳这样以氢为燃料的稳定恒星，维持生命所必需的环境，就需要让恒星花费数十亿年来完成炼金术的过程。这就是为什么我们不应为宇宙如此古老而感到惊讶的原因。任何化学复杂性所需的砖块都需要经过长时间的烧制才能产生。又因为宇宙在膨胀，如果宇宙已经很老了，那么宇宙的体积必然会很大——百亿光年之广。假如宇宙跟银河系一样大，让其一千亿颗恒星孕育行星系统的话，看起来有足够大的空间滋养大量的生命。但这样一款经济型宇宙的年龄也就一个月左右。这点时间根本不足以让恒星演化，烧制出生命复杂性所需的砖块。

如果把目光转向可见宇宙的其他方面，我们会发现，为了允许生命存在，这些方面的属性也必须落入相应的取值范围。我们已经提到过，宇宙的膨胀速度可以比临界速度更快，或更慢。如果宇宙年轻时的膨胀速度过快，恒星和星系就无法形成了，因为物质会分得太开，牛顿、金斯和栗弗席兹的引力不稳定性就无法让物质结团了。相反，如果膨胀速度比临界速度还慢很多，物质在宇宙早期就会结成致密的团块，宇宙中就会充满黑洞而不是以氢为燃料的恒星。我们再一次发现，不该为宇宙的膨胀速度非常接近临界速度而感到惊讶。如果不接近的话，我们就不会存在于宇宙中了。

为什么宇宙的膨胀速度在临界速度附近保持了如此长的时间？暴胀宇宙模型提供了一个简单的解释（图10.2）。[4]乍看之下，这种思维方式很怪异。我们通常称之为“人择原理”。这不是一个理论，无法被证明

或证伪（它是正确的），而是一种方法论的原理，能够防止你从证据中得出错误的结论。这个原理告诉我们，宇宙拥有生命的存在和演化所必需的属性，所以我们不可能发现自己生活在一个没有这样属性的宇宙中。许多事情都可以归结为，烧制生命复杂性的单元需要足够长的时间，这些结论我们已经知道了，所以可见的宇宙必然非常宽广。这同时也说明宇宙必然十分寒冷、十分昏暗。经过了百亿年的膨胀，空间中辐射物质的温度必然持续下降，宇宙的尺寸每增大一倍，温度就下降为原来的一半。漫长的膨胀降低了宇宙的物质密度，所以今天平均每立方米的空间中只有一个原子。经过了百亿年的膨胀，物质的密度如此之小，辐射的温度如此之低，宇宙中几乎没有什么东西能够照亮我们的夜空。即使利用爱因斯坦的 $E=mc^2$ 大法将宇宙中所有物质瞬时转化为光，我们也看不到任何显著的现象。空间中辐射物质的温度只是经历了一个不起眼的上升，从3开上升到15开。因此，夜空十分昏暗是因为宇宙十分古老。宇宙的夜空曾经十分明亮，到处都像太阳表面一样耀眼夺目，那时宇宙大约25万岁，尺寸只有现在的千分之一左右。当时的温度太高了，任何恒星、行星、分子、原子都无法形成。没有观测者能见证如此夺目的夜空。

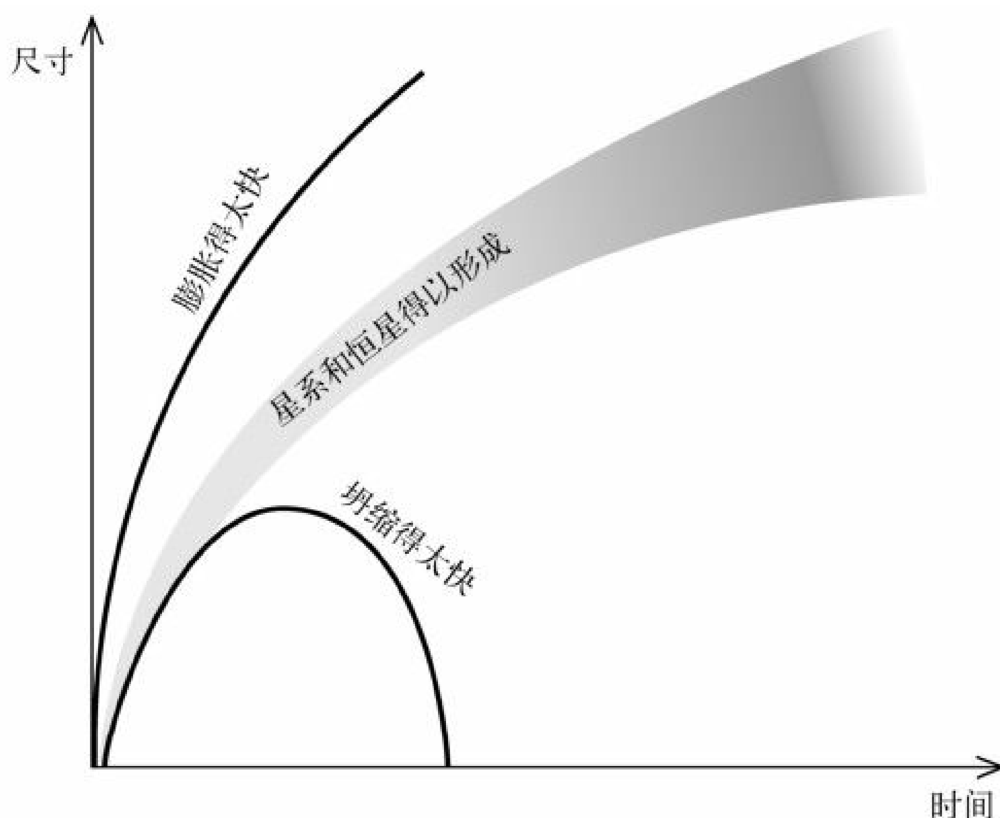


图10.2 膨胀速度偏离了临界速度的宇宙或者会在恒星形成以前就经历一

场大塌缩，或者膨胀得太快，物质抵不过宇宙膨胀的影响，无法结合成星系和恒星

这种看待我们的宇宙结构以及其他我们能想象到的宇宙结构的新视角，统一了我们对生命和宇宙的认识。这种视角同时也能防止我们从天文观测证据中得出错误的结论。在第4章中我们看到，保罗·狄拉克发现不同自然常数的取值之间存在某种巧合，如果将特定的常数组合在一起，就能得到两个约等于10^{39}的大数。他认为，这两个数值如此之大，又如此接近，说明自然常数的背后必然存在某种未知的方程或者未知的自然法则。正如我们前面所看到的那样，狄拉克的想法的新奇之处在于，其中一个大数来自于真正的自然常数的组合，而另一个大数则涉及宇宙的年龄，狄拉克把宇宙的年龄也当成了某种常数。但是宇宙的年龄会随着时间增长，为了让两个10^{39}保持相等，他就得强行要求其他传统的自然常数也要随着时间变化。他选择了牛顿的万有引力常数G，要求它会随着宇宙年龄的增长而变小，与时间呈反比。这是一种激进的想法，但这个想法没什么根据。罗伯特·迪克指出，大数之间所谓的巧合不过是换个说法说明，我们观察宇宙时，宇宙所经历的时间刚好来得及让恒星形成，并处于稳定地燃烧氢元素的时期。我们不可能在恒星形成之前就开始观察宇宙，观测者也不可能在所有的恒星都耗尽了各自的能源后还能悠然自得，所以狄拉克说的巧合一点儿也不巧妙。由于忽略了天文观测中观测者所扮演的角色，狄拉克就得到了错误的结论。诞生于大爆炸的宇宙存在几个特殊的时期，如恒星可以形成并能够稳定地存在的时期，原子可以形成的时期，等等。稳态宇宙理论给宇宙套上了一套紧身衣，要求其历史中不存在特殊的时期，但这个要求严重地违背了天文学的观测证据。

狄拉克并没有公开接受迪克的批评，而是声称，他很乐意承认我们不会在恒星形成之前就开始观察宇宙，但他又认为，说恒星耗尽能源后就不可能存在观测者是没有道理的：

按照迪克的假设，只有在一段特殊的时期内才有可能存在宜居的行星。在我的假设中（也就是G会随时间变化），在无穷远的未来，这种行星依然能够存在，而生命永远不会灭绝。这两种假设谁对谁错尚未有定论。我倾向于那个允许生命无限延续的假设。[5]

值得注意的是，早在1933年1月，狄拉克就已经确定了自己的生命观，并在笔记本上花了三页的篇幅阐述了他的理解，这些文字和上述生命在宇宙中无限延续的想法极其相似。他的哥哥菲利克斯自杀后不久，狄拉克拒绝用传统的宗教方式进行悼念，而是转向另外一种形式的信仰，即为生命的存在寻找意义和目的，并将这些想法记了下来。他的私人笔记已经隐隐透露出他后来的宇宙观：

［我的］信仰是，人类会永远繁衍下去，能够毫无限度地发展和进步下去。为了我内心的安宁，这是我必须作出的假设。如果一个人能够以某种方式为无限的发展环节贡献绵薄之

力，活着就有价值的。[6]

虽然在公开场合对观测者决定宇宙学观测的想法不置可否，狄拉克在与伽莫夫私下联系时还是有所表示的。正如我们所看到的那样，如果G 的数值曾经比现在大得多，就会改变太阳的演化和地球的温度，导致非常严重的后果，于是狄拉克的引力常数随时间变化的理论很快就被否定了。但是狄拉克又提出了一种很不靠谱的说法为之辩护，假设太阳系围绕银河系转动时，碰巧路过了巨大的尘埃云。太阳额外地吸积了一些物质，质量有所增加，刚好抵消了G的数值变小所产生的影响，于是太阳对地球的引力能够保持不变。伽莫夫认为这相当不太可能，而且很“不雅观”——伽莫夫所说的“雅观”，意思是狄拉克应该用优美的数学形式来表达这个神话故事般的物理理论。为什么整个宇宙的G都变化了，而太阳吸积的物质正好能够抵消G的变小？然而，在1967年11月写给伽莫夫的信中，狄拉克竟然出人意料地用某种人择原理的说辞为自己进行了辩解：

> 我不认同你对我的吸积假说的反对理由。我们可以假设，太阳曾经穿过某种致密的云层，致密到可以让太阳俘获充足的物质，以保证地球在 109年中都保持适宜的温度。你可能会说这不太可能，因为云层的密度不太可能刚好达到这个目的。我同意。这不太可能。但这种类型的不太可能并不是重点所在。假如我们考虑所有存在行星的恒星，其中只有极个别的恒星穿过了密度适当的云层，于是相应的行星才得以长时间保持恒定的温度，以使高级生命能够发展壮大。正如我们先前所想的那样，存在人类的行星不会很多。但是只要存在一个这样的行星就足以解释现实了。因此，太阳可能有那么一段极不寻常又不太可能发生的历史，我们没有理由反对这个假设。

有趣的是，狄拉克从来没有用这样的理由来解释大数巧合。

回想一下我们会发现，这种类型的“人择”理由同样也使得20世纪50年代的稳态宇宙模型看起来疑点重重。在大爆炸理论中，宇宙的年龄大致与膨胀速率的倒数相当，天文学家把宇宙当下的膨胀速率观测值称为“哈勃常数”。

在稳态宇宙理论中，宇宙不存在年龄一说（年龄是无限大的），因此，膨胀速率就成为宇宙的一个完全独立的属性，需要单独进行解释。天文观测表明，像太阳这样的稳定恒星的年龄非常接近宇宙的年龄，但显然比后者小一些，这个特点在大爆炸模型中是完全正常的。形成星系，然后形成恒星，然后是行星，然后是天文学家，历史环环相扣。[7]因此，在大爆炸理论中，宇宙现在的膨胀速率大约等于恒星年龄的倒数一点儿也不奇怪。[8]但在稳态宇宙理论中，这完全是个巧合。

天文学家逐渐搞清楚了宇宙中生命的存在与否对各个物理学常数的敏感度。就像宇宙的膨胀速率稍微一变，就会对生命产生剧烈的影响一样，改变自然界中相互作用的强度，改变基本粒子的质量，都会导致恒星和原子无法形成，进而改变宇宙的历史进程。[9]这种关于生命对物理学常数和宇宙结构的敏感度（或不敏感度）的计算，后来被称为“人

择”观点。[10]这种观点有时会说，我们观测到的宇宙在某些方面被“微调”过了，因此有利于生命的演化。如果某些常数的取值发生一点点改变，那么幸运之窗就会关闭，原子和恒星就无法形成了，生命的复杂性也就无法显现了。

确定其中一些概念的含义是一个非常棘手的问题。我们定义的“生命”包含多广的范围？自然常数的“一点点”变化又是多少？所有这些常数都是相互独立的吗，还是说这只不过是一种表面现象，实际的原因是我们缺乏一个完备而统一的物理学理论？

直到20世纪80年代末，这种看待现实的观点都还是让人觉得古怪。大多数宇宙学家认为，我们只有一个宇宙，宇宙拥有它所拥有的性质；从科学的角度讲，再没什么可说的了。你可以走得稍微远一些，设想在形而上学的意义上还存在很多（甚至所有）可能的宇宙，而我们的宇宙则恰好处于这个展览馆里一个允许生命存在的区域。[11]为了解释这个状况，你可以借助哲学或宗教的观点来讨论宇宙是否允许存在生命。但假如宇宙只有一个，而它又不适宜生命存在的话，我们就不可能坐在这儿讨论这个问题了。

只有一个宇宙的论调还有一个关键的假设，那就是所有自然常数以及宇宙的所有性质都是完备而唯一的，其中不存在任何灵活性，不存在另一个与我们的自然法则、自然常数略有不同的宇宙。想象你用麦卡诺组合玩具搭成了一个三角形和一个正方形，并在拐角处拧上螺丝。这两者大不相同。正方形不是刚性的：如果把顶部向右拉，把底部向左拽，就会将正方形流畅地变成一个平行四边形（图10.3）。但是三角形可不吃这一套，它是刚性的。[12]

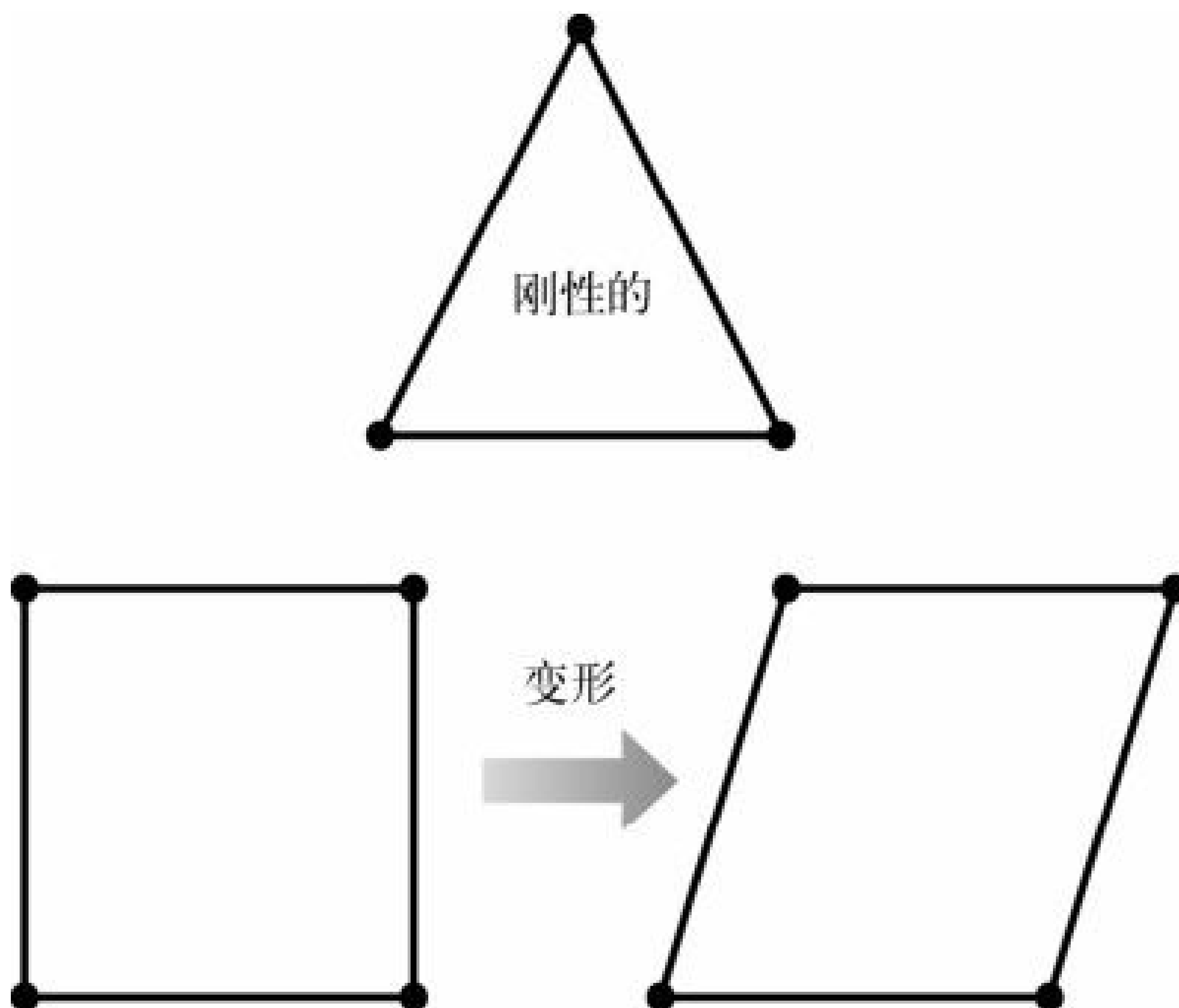


图10.3 铰链将杆子连接成刚性的三角形：如果不把它折弯，就无法将它自然地变成别的形状。铰链连接成的正方形就不是刚性的了：我们可以将它自然地变成平行四边形

物理学的定律和常数是像刚性的三角形一样，独一无二；还是像灵活的正方形一样，存在无限种可能，其中一些比较类似，而另一些则迥然不同呢？实际上，在正方形的例子中，我们已经对其连接处和边长施加了某种限制。同理，即使存在多套不同的常数和定律，也不意味着一切皆有可能。在可量化的自然常数背后，仍然可能存在一种支配一切的约束条件。像爱因斯坦这样寻找一个万有理论（他的“统一场论”，或者爱丁顿的“基本理论”）的物理学先行者，坚信物理学家能够找到关于这个世界的唯一的终极描述。事实上，他们试图用数学的对称性和纯粹的思考来寻找这样的基本原理。这个目标远远超出了当时的实验水平。你只能寄希望于什么时候有一种美妙的数学结构会从书中跳出来，然后说

道："这就是唯一的可能！"

如果你问爱因斯坦如何看待其他宇宙，或者其他类型的多重宇宙，其中的自然常数的大小与我们的观测结果大相径庭，他可能会不怎么感兴趣。有一次，爱因斯坦给他的老朋友、一辈子的通信对象伊尔斯·罗森塔尔–施耐德（Ilse Rosenthal-Schneider）写信说道：

> 在一个合理的理论中，不应该存在只能凭经验确定的无量纲常数。当然，我无法证明这一点。但是我也无法想象在一个统一的、合理的理论中，会存在这样一种常数——如果造物主心血来潮对此略加改变，一个性质不同但同样有效的世界就会出现……从纯粹逻辑的角度讲，如果自然法则中的无量纲常数可以取别的值，那么这样的常数就不应该存在。在我看来，凭借我对"上帝的信仰"，这一点是显而易见的，但是持有相同观点的人寥寥无几。[13]

可能的宇宙

> 曾经有一段时期人们认为，一切都会发生改变，但实际上一切照旧，只不过换了种形式。
>
> ——亨宁·曼凯尔（瑞典侦探小说作家）[14]

自1990年起，旧的观念逐渐发生了转变，人们不再把宇宙和其中的常数和定律看作刚性的三角形。宇宙的很多性质曾经被看作硬道理，如今却被解释为某种深层原理所导致的对称性破缺。如果你将一块铁棒从极高的温度冷却到 770 摄氏度以下，那么铁棒就会变成磁铁。铁棒的哪一头变为北磁极，哪一头变为南磁极，几率是相等的。在铁棒变冷之前，你没有办法预测哪一头会变成北磁极；当铁棒的温度跌至770摄氏度以下时，原子的对称性就破缺了，铁棒的南北磁极就随机地产生了。

这类随机的"对称性破缺"过程能够决定宇宙的许多重要性质——正反物质不对称性、原子物质的密度以及宇宙级磁场的强度和方向。在这个前提下，人择原理的概念就变成了另外一个样子：如果宇宙不同角落中的对称性以不同的方式破缺，就会产生性质各不相同的新宇宙。

混沌的永恒暴胀宇宙存在的一些独特行为说明，引入人择原理至关重要。永恒暴胀宇宙的很多角落比我们的可见宇宙还宽广，于是我们得计算一下，产生一个拥有某种特定性质的宇宙的概率有多大。如果引入对称性破缺所导致的复杂性，我们就会发现，永恒暴胀所孕育的众多"宇宙"也可以表现出截然不同的物理性质。不同角落的自然常数可以有不同的取值；弦理论预言，多重宇宙不同角落的巨大空间维度（甚至时间维度）也会各不相同；基本的相互作用数目和种类也可能各不相同，其中每一种组合都体现了弦景观的一种真空态。

在传统的科学方法看来，这个问题实在太艰巨了。人择原理教育我们，研究不同宇宙从多重宇宙中产生的概率时，我们应该集中精力研究一些特殊的宇宙，在这些宇宙的某个演化时期，能够产生复杂性、生命以及存在意识的观测者。这个子集也包括我们的可见宇宙，但是我们并不想把生命定义得过于保守，不然就没法产生跟我们的意识类型截然不同的生命了。目前，想要确定宇宙中能产生观测者的条件实在太难了。

我们甚至连定义“生命”所需的必要条件都不清楚。我们所了解的只不过是一些弱的充分条件，这些条件差不多都来自于我们对自身的认识。

这导致我们很难预言，多重宇宙的众多可能性中哪一种宇宙最有可能产生生命。为了回答这个问题，我们得搞清楚哪些物理定律、宇宙的哪些结构触及了产生生命复杂性的关键。我们只了解其中一些关键性因素，但能肯定的是，真正的约束条件一定比我们目前所认为的要宽松得多。这是因为我们还没找到能将自然界四种基本相互作用（电磁、弱、强和引力）统一起来的万有理论。当我们找到这个理论时，也只不过了解了更多决定基本相互作用特性的自然常数之间的关联性和依赖性。目前，我们只知道改变放射性物质的衰变率时并不用担心这会对引力或原子的结构产生什么影响。一个完备的统一理论应该表现出所有这样的关联性，并且能够保证，其中的物理规律应该是牵一发而动全身的。

如果我们学会了计算这种概率和条件概率的方法（我确信这要不了多久就能学会），我们还得面对另一个极为复杂的问题，也就是研究“观测者”对物理规律的依赖性，然后用一个完备的理论来描述这些依赖性。用这种方法我们也许就能找出几种不同的宇宙族，其中的每一个宇宙都能满足生命产生的必要条件。接着，我们可以研究哪一个宇宙才是最可能出现的。但是假如最可能的那个族中的宇宙跟我们看到的宇宙不太像呢？我们是会断定理论的某些部分（或者全部）错了，还是会断定我们的宇宙并不是其中最可能的宇宙呢？假如我们找到的几个族大体上看都差不多的话，那么我们又该如何进行抉择呢？

这些问题都没有答案，所以我们必须把人择原理铭记在心。无论多重宇宙中各种宇宙出现的概率有多大，最后都必然要考虑到观测者的存在。我们不得不承认，在上述讨论中，我们自己（以及其他能收集和处理信息的实体）也是我们想要解决的问题的一部分。

自制的宇宙

哲学的精髓在于，从一些极其简单的、没什么可说的概念出发，得出某些令人难以置信的悖论。

——伯特兰·罗素（1872～1970）[15]

不断自我繁殖的宇宙尝试了所有可能的真空态，以及各式各样的物理常数、空间维度和基本相互作用的组合模式，这是我们得到的最接近“所有可能世界”的图景。从前的哲学家经常从形而上学的角度思考“所有可能世界”。关于我们所生活的世界是否是“所有可能世界中最好的一个”，哲学家们从各个角度进行了激烈的争论。[16]然而，暴胀宇宙并没有尝试宇宙所有属性的所有可能变化，因此，它并没有创造出形而上学意义上的“所有”可能世界，而只是尝试了弦理论允许的所有自洽物理规律的版本。这些版本的数量十分惊人，不过是有限的，约为10^{500}。

这个繁殖过程由物理定律所决定。它并不像中世纪基督神学中的宇

宙创生，是一种神秘的“从无中创造”。乍看之下，即使在这样的限定条件下，人们依然会觉得将某种东西“创造”出来是不可能的。但这并没有违反任何物理学守恒律。设想一次自发的量子涨落产生了一个虚的粒子和一个虚的反粒子。如果每个粒子的质量都是M，就要求总能量为 $2Mc^2$。但如果这些粒子之间存在某种吸引的相互作用，例如引力或电磁力（如果它们电性相反的话），那么这种相互作用就会贡献一种负的“势能”，因为这种势能一旦被释放出来，就能转化为粒子的运动。这就像你扔出一块石头，在地球引力的作用下，石头会掉落在地面上。

值得注意的是，$2Mc^2$ 的正的能量和吸引作用产生的负的势能加起来可以恰巧为零，于是从量子真空中“创造”两个粒子并不消耗任何能量（图10.4）。

这样的过程，再加上“宇宙”在永恒暴胀之中自发繁殖的过程，使得宇宙学家开始设想，我们能不能人工地激发这种繁殖过程呢？[17]我们能不能通过激发某一次涨落，产生像永恒暴胀一样的效果，从而“创造”出一个宇宙来呢？至少从原则上来讲，物理学定律是否允许这样做呢？

这个问题没有确定的答案，有的人证明不可能，有的人证明可能，然而其中似乎也产生了极其不必要的副产品，如无穷大的密度。与此同时，也有人，如亚利桑那大学的爱德华·哈里森，认为这开辟了高级文明得以操纵宇宙的前景。[18]

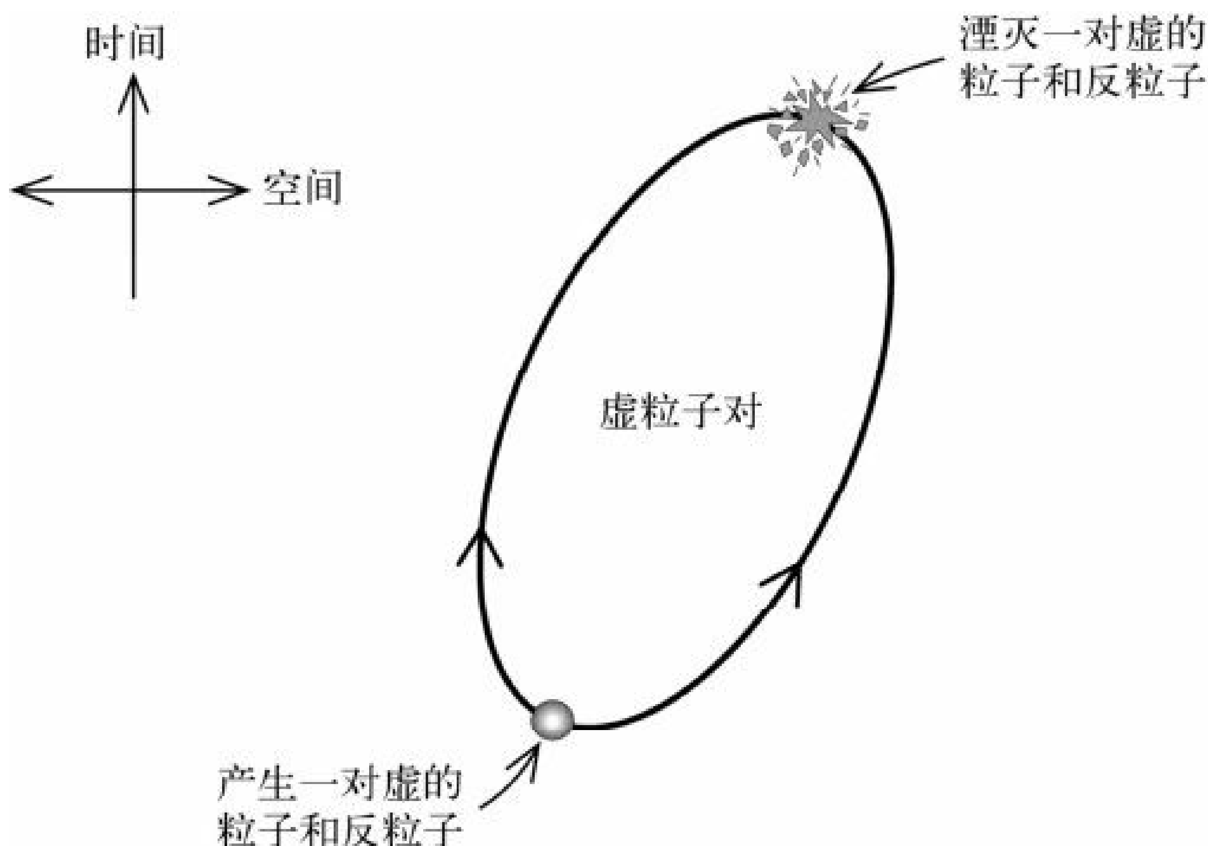


图10.4 产生一对虚粒子（虚的粒子和反粒子），然后湮灭

设想宇宙中存在一些高度发达的文明，他们已经完全掌握了在附近空间制造特定量子涨落的方法，这些涨落会迅速引发暴胀，产生新的婴儿宇宙。那些超级宇宙学家也已经完全掌握了自然常数和物理定律的奥秘，知道什么样的组合才会允许生命的产生。假设（考虑到可能性的数量以及从中进行选择的计算复杂性，这是一个非常大胆的假设）他们有能力控制对称性破缺的过程，能通过控制温度下降来选取特定的真空态，那么他们就能够强行孵化出新的宇宙，其中的自然常数和物理定律比他们自己的宇宙还要适合生命的繁衍。加快自我繁殖的速度，快进到好几代宇宙之后，更加发达的文明就会产生（而且产生得更加容易，因为新宇宙出现生命的“几率”被上一层文明精心调整过了），他们于是拥有更加先进的技术来调制更新的宇宙。哈里森设想，我们的宇宙之所以如此恰到好处，如此适合生命的繁衍和演化，或许正是因为宇宙经历过这种类型的强制孵化[19]，它的性质都被微调过了。

这是个异想天开的想法，但其中有一点值得一提。一旦宇宙中演化出能够操纵宏观环境的智慧生物，宇宙学就成了一门不可预知的学科，就像经济学和社会学一样，因为人类的行为不但从实践上无法预测，从

理论上也是无法预测的。

自然选择的宇宙

有一种学说认为，不管什么原因，如果有人发现了宇宙的本质和目的，宇宙就会立刻消失，取而代之以某种更加离奇、更加说不清道不明的东西。还有一种学说认为，这种事已经发生过了。

——道格拉斯·亚当斯（1952～2001）[20]

另一个尝试解释自然常数特殊取值之谜的人是美国物理学家李·斯莫林。他不像哈里森那样异想天开，没有将智能观测者的行为也考虑在内，让他们操纵自己制造的宇宙来产生更加发达的文明。斯莫林考虑的是托尔曼想法的一种变体，即一种不断振荡的、能从收缩反弹为膨胀状态的闭合宇宙。[21]斯莫林打算把这种想法应用到黑洞内部的引力塌缩中去。毕竟，如果你在一个黑洞的内部，你就会发现它实际上跟一个闭合的宇宙没什么区别。如果你从中心出发向外走，永远也不可能穿过它的边界（图10.5）。你会发现自己总是被拉回到中心的奇点处，也就是你出发的地方。这就像一个闭合的宇宙从奇点开始往外膨胀，到最大值后再向奇点收缩。

第一个想法是，每当有黑洞形成，物质纷纷落入奇点的时候，新的宇宙就会从奇点中“跳”出来。第二个想法跟约翰·惠勒在1957年时提出的想法一样，振荡的宇宙每进行一次反弹，自然常数的大小就可能发生变化——惠勒称之为“再加工”。[22]斯莫林假设，每当有黑洞形成，其中的物质无情地被奇点吞没时，这种反弹就会发生，在新产生的膨胀中，自然常数在原有的基础上进行了微小的变动。结果就形成了一种自我繁殖的局面，随着每一次繁殖的发生，其中的常数都会进行微小的变动。[23]

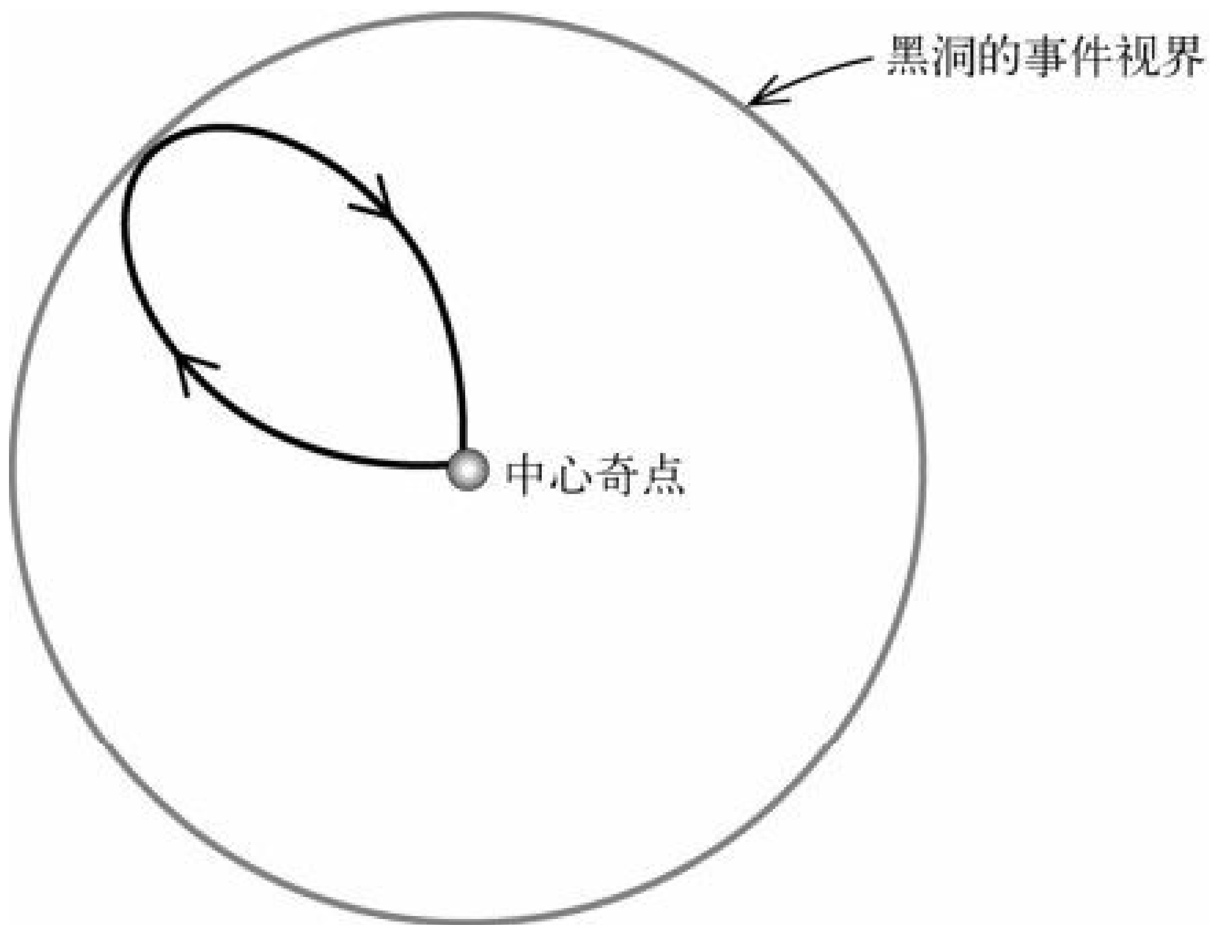


图10.5 一个粒子从黑洞的中心奇点发出，当它达到黑洞的视界时，距中心最远，然后它必然会落回到中心奇点处。它的轨迹很像闭合宇宙中的一个粒子，从大爆炸开始膨胀，然后又开始坍缩，经历一次大塌缩

如果我们让这个过程持续很多代，又会如何呢？宇宙中形成黑洞，然后每一个黑洞都会导致更多宇宙的形成，每一个宇宙的常数都略有变动。从这个过程的长期行为来看，按照他们的说法，我们会发现自己身处一个特殊的宇宙，这个宇宙的常数会导致黑洞产生率达到最大，因为这种宇宙的数量最多，我们存在其中的概率也最高。这跟生物学中的演化差不多。在生物学中，如果有一种可遗传的属性能导致子代数量的最大化，那么拥有这种属性的生物就会比其他生物更加普遍地存在。在这个理论中，能产生更多黑洞的引力属性就让这样的宇宙遍地开花。

如果住在这样一个历经了许多代引力的自然选择过程，最终还能脱颖而出的宇宙中，也许你就能预言自己应该会看到什么现象。因为演化选择了黑洞产生率最大的那个宇宙，所以我们宇宙里的自然常数估计已经使黑洞产生率达到或差不多达到了局部的极大值。

这就意味着如果我们的常数发生微小的变动，黑洞产生率就会下降。这听起来不错。斯莫林提议，我们可以做一些思想实验，看看略微修改物理学常数会导致什么样的后果。如果这些改动导致黑洞产生率上升了，那我们能因此否定这个理论吗？斯莫林提出这套理论的目的是想在不借助人择原理或者多重宇宙概念（在多重宇宙中，几乎所有的可能性都实现了）的情况下，解释自然常数的取值为什么很特别。

可惜，这个简单的想法并不可行。且不说是否存在微小的变动导致黑洞产生率增加的问题，其中还有两个严重的问题。尽管这个理论的目标之一是将宇宙学解放出来，使我们不再苦恼于人类对理论推演的影响，但我们不难发现，这个理论本身就是人择原理重要性的最佳诠释。我们可以料想我们宇宙的黑洞产生率已经达到了最大，并且这样的宇宙允许观测者存在。如果黑洞含量最多的宇宙只包含辐射和尺寸非常小的黑洞，那么其中就不会存在任何智慧生物。为了得到该理论的一个可检验的结论，你得知道所有自然常数的取值范围，以确保观测者的存在——但这个取值范围不一定对应最可能的取值，甚至可以相差甚远。于是，你很可能会说这个理论预言了一组自然常数的取值，在允许生命存在的取值范围中能使黑洞产生率达到最大。这恰好印证了我们在图10.1中所强调的内容。

该理论的第二个问题更加严重。它假设，当你调整自然常数的取值时，黑洞产生率存在局部最大值。这是一个非常强的假设，仅对我们已知的一个常数来说，这个假设就不成立。如果我们略微改变牛顿常数 G的大小，从而改变了引力的强度，那么我们就会发现结果在局部是单调变化的。由于热力学第二定律的影响，自然常数每变动一次，并非只可能导致黑洞的数量增加，还可能导致G的数值越来越大。此时，宇宙的引力就如长江后浪推前浪，一代更比一代强。[24]

虚拟的宇宙

> 在其他条件相同的情况下，如果你是生活在一个虚拟世界，那么你会更不在意别人，更愿意活在当下，会让你的世界看起来更容易变得富有，并希望试图更多地参与重要事情，会让自己更富有娱乐精神，更值得称道，以及让名人乐于围着你团团转。
>
> ——罗宾·汉森（美国经济学家）[25]

一旦你认真对待所有（或几乎所有）可能的宇宙都可以（或确实）存在的想法，那么一种滑坡谬误（slippery slope）①就会出现在你的面前。我们已经看到，一个文明的科技水平只要比我们发达一点点，就有能力进行一场宇宙模拟实验，其中可以产生多种能够相互交流的智慧生物。[26]这个文明会拥有强大的计算机，其计算能力要比我们的计算机高一个巨大的量级。他们不但像我们这样，能够模拟天气变化或星系的形成，还能进一步看到恒星和行星系统的形成，于是通过对虚拟世界的研究，他们的行星地质学和地形学也得到了发展。接着，将生物化学的法

则和天文学的模拟结合起来以后，他们就能看到生命和意识的演化过程（根据他们的需要，这些过程都可以被任意地加速）。正如我们能观察果蝇的生命周期，他们也能跟踪生命的演化进程，观察文明如何发展，如何相互交流。他们甚至可以围观虚拟生物之间的争论：天上是否存在一个“大程序员”，他不但创造了宇宙，还可以随意挑战他们已经习惯遵从的自然法则。

① 武断地将某个可能性引申成为必然性，然后串联这些不合理的因果关系，推断出一件毫无关联的结果，这就是滑坡谬误。——译者注

一旦有谁掌握了模拟宇宙的能力，虚拟的宇宙就会如雨后春笋般冒出来，很快就会超过真实宇宙的数量。而且尼克·博斯特罗姆断言[27]，如果随机选出一个智慧生物作为考察对象，那它更可能生活在虚拟现实之中，而不是真正的现实之中。[28]

这个惊人的结论引发了更多讨论，如果我们极有可能生活在虚拟现实之中，自己只不过是其中的虚拟生物，那么我们该怎么办才好呢？正如在本节开头的引语中所说的，罗宾·汉森认为，你应该做的就是努力创造机会延长自己在这场模拟中的存活时间，或者寻求将来再被模拟一次，尽管这个忠告听起来像是很多人都信奉的个人奋斗策略，而不论他们认为自己生活在哪种类型的宇宙中。对此，保罗·戴维斯回应道，我们极有可能生活在虚拟现实之中，这以反证法证明了涵盖所有可能性的多重宇宙的存在——但这无疑会让我们从宇宙中获取可靠知识的愿望日渐渺茫。[29]

有些人认为宇宙是由一个“大设计师”专门为生命的存在而设计的。为了反对这种论调，一些宇宙学家明确地提到了我们所介绍过的多重宇宙学说以及暴胀宇宙理论。[30]无论生命所需的条件多么苛刻，如果在无限的空间中存在一切可能性，多重宇宙中就必然会存在一个宜居的宇宙。但在讨论中常常会被忽略的一个关键问题是，在所有的可能性中，宜居地带所占的可能性究竟有多大？没人知道答案。因此，那些人认为，多重宇宙只不过是一种逃避微调问题的简便方法。

对于多重宇宙理论的这两种观点都希望能够避免走向形而上学。但可惜多重宇宙的事情没有这么简单。我们已经看到，一旦有意识的观测者能够随意干涉宇宙的进程，而不再是被动的“观测者”，那么结果就会是再次出现数不尽的神，他们以模拟者的面貌出现，掌管着手中虚拟现实的生杀大权。模拟者能确定虚拟世界所受的各种规律，同时也能对之作出修改。他们能设计出一种人择原理的微调方案。他们可以随时拔掉插头，也可以参与到模拟之中或者置身事外。他们可以冷眼旁观，看着

虚拟生物在讨论是否存在一个掌控全局的神。他们也可以制造奇迹，或者将他们的道德准则引入到虚拟现实之中。他们永远不用担心自己会造成什么伤害，因为他们的玩具现实不是真的，不是吗？他们甚至可以观察到虚拟现实里的文明发展出了极高的科技水平，于是在虚拟现实中再模拟出一层虚拟现实出来。

面对这种错综复杂的情况，我们有可能将虚拟的现实与真正的现实区分开来吗？如果我们只不过是在虚拟现实当中进行科学观察，我们又能预期观察到什么呢？

首先，模拟者会避免在虚拟世界中运用整套自洽的自然法则，因为这会比较复杂，而且他们大可只需拼凑出接近"真实"的效果。当迪斯尼的动画片要表现湖面的反光时，他们不会利用量子电动力学和量子光学的定律去计算光的散射，因为这得动用超强的计算能力。相反，用一些合理的经验法则就能模拟光的散射过程。这比真实的散射过程简单许多，但却能营造出真实的感觉——只要人们不要离得太近就行了。如果模拟者只不过想玩玩而已，选择这样经济实用的方案来模拟现实就很有必要。但如果模拟程序之中真的包含这样的不足之处，就可能偶尔会出现穿帮镜头，也就是业余剧团表演话剧时常见的"吱吱作响的布景问题"，而且也许这种穿帮镜头在虚拟世界内部也能看到。

即使模拟者一丝不苟地模拟了自然法则，他们也无法超出自己的能力范围。虽然他们对自然法则的认识十分深刻，这种认识也不可能是完备的（一些科学哲学的研究者会论证说，事实绝对会是如此）。他们也许知道很多模拟宇宙所需的编程知识，但对于自然法则的了解总会存在一些空白，或更糟地，存在错误之处。这些漏洞应该是难以捉摸、不易发现的，不然我们设想的"发达"文明就配不上"发达"二字了。

这些缺点不会妨碍模拟实验开机并长时间地流畅运行，但慢慢地会有一些缺陷逐渐显现出来。最终，这些缺陷产生的效应如同一场雪崩，导致虚拟现实"死机了"。唯一的解决方案就是让程序员每发现一个缺陷，就立刻打上一个补丁。家里有电脑的人早已对这种解决方案习以为常了。我们会定期收到系统的更新提示，以便修复程序设计者原先没有预料到的漏洞，或者防止新型电脑病毒入侵。模拟开始以后，模拟者也会采取这样的临时防护措施，更新自然法则的版本，并向其中添加一些他们新学会的东西。

在这种情况下，逻辑矛盾在所难免，模拟的法则时不时地就会被破坏一次。虚拟世界中的居民（尤其是虚拟科学家们）偶尔会被他们得到的实验结果搞糊涂。而虚拟天文学家们则可能，打个比方说，从观测中发现他们所谓的自然常数正在缓慢地变化着。[31]

统治虚拟世界的法则甚至也可能突然失灵，这很可能是因为模拟者

在应用一种在模拟其他复杂系统时已经证明行之有效的手段，即使用纠错代码，让事情都回到正轨。

以我们的遗传代码为例，如果任由这些代码运行，人类就延续不了多长时间。误差日积月累，必然会产生致命的危害。幸好存在一种纠错机制，能够识别和修正遗传代码中的错误，使我们免受退化的威胁。许多复杂的计算机系统中都有这种内置的“拼写检查”功能，以防止错误的积累。这种纠错码是由贝尔实验室的理查德·海明（Richard Hamming）在1950年最早提出的。

如果模拟者利用了纠错码防止整个模拟（以及更小尺度上的我们的遗传代码）出错，那么虚拟世界的运行状态和法则就会时不时地被修正一下。这样，神秘的改变就会突然发生，违背了虚拟科学家通常观测和预测的自然法则。

我们还可以预期，虚拟现实的各处都具有最高程度的计算复杂度。虚拟生物的计算复杂度应该与最复杂的无生命系统（如天气系统）的相当，借用史蒂芬·沃尔夫拉姆的说法（虽然他是基于完全不同的原因而提出的）就是“计算等价原理”。[32]

关于从内部区分虚拟和现实的问题，一个最普遍的担心就是，模拟者能够隐隐预感到不妙，因而会预先调整模拟的过程，防止故障的发生。这套新的虚拟现实可能又会与真正现实产生新的不一致，但他们又可以通过再次预判而防止故障的发生。

问题在于，把其推至极致时，这种先见之明是否可能。这个问题很像卡尔·波普尔所考虑的计算机的自指能力极限（self-referential limits）。[33]最近，唐纳德·麦凯又在另一种背景下利用了同样的论证。麦凯借此否认存在这样的逻辑可能性，即如果事先将一个人的命运告诉他，他未来的行为仍然是可预测的。[34]

只有不把预测结果告诉你时，关于你未来一切活动的预测才可能是正确的。一旦我的预测结果被你得知，你就有可能改变未来的计划。[35]因此，关于你未来的活动，不可能作出没有条件限制的预测。世俗事物中也会出现这样的例子，例如预测选举的结果，对选举结果的公开预测不可能不考虑预测本身对选民的影响。[36]这种类型的不确定性在原则上是不可约的（irreducible）。但如果不公开预测结果的话，其正确率可以达到100%。

这一切都说明，如果我们生活在虚拟现实中，我们就会偶尔遇到自然法则的突然失灵，会发现自然法则和自然常数随着时间的推移而发生了微小的变动。[37]这也让我们第一次意识到，对我们理解真正的现实来说，自然的缺陷就同自然法则一样重要。

这些略有跑题的讨论已经进入了通常由科幻作家占据的地盘，但我

只是想指出，如果我们认真对待存在无数个可能的世界（它们穷尽了所有的可能性）的想法，那么也许会得出许多异乎寻常的推论。我们可以设想，在将来，现有某些科学和技术的进步可以让我们的后人对这些推论进行检验。我们甚至不用发明新的科学，技术进步本身就可以把这搞定。[38]我们所见宇宙的本质及其可能的虚幻性让我们触目惊心，甚至令人忧心忡忡，这不禁让我们想起了英国哲学家大卫·休谟（1711～1776）在18世纪末时所说的话。

休谟在《自然宗教对话录》中驳斥了许多在当时非常流行的关于上帝存在性的论证，将矛头直指这些论证当中对于神创论完美本质以及神的唯一性等假设。[39]这里摘录了休谟关于“多重世界”及其可能存在的缺陷所作出的讨论①：

① 摘自陈修斋、曹棉之的译文（［英］休谟，《自然宗教对话录》，北京：商务印书馆，1962年，页39，页41）。——译者注

你必须承认，从我们狭小的观点，我们是不可能说出，这个宇宙体系比起其他可能的、甚至真实的宇宙体系来，是否包含有任何巨大的错失，或值得承当任何巨大的赞美……假如我们考查一只船，对于那个制造如此复杂、有用、而美观的船的木匠的智巧，必然会有何等赞叹的意思？而当我们发现他原来只是一个愚笨的工匠，只是模仿其他工匠，照抄一种技术，而这种技术在长时期之内，经过许多的试验、错误、纠正、研究和争辩，逐渐才被改进的，我们必然又会何等惊异？在这个世界构成之前，可能有许多的世界在永恒之中经过了缀补和修改；耗费了许多劳力；做过了许多没有结果的试验；而在无限的年代里，世界构成的技术缓慢而不断地在进步……

就他所知道的，这个世界，比起一个高的标准来，是非常错失，非常不完全的；这个世界只是某个幼稚的神的初次的尝试作品，后来他抛弃了它，并对于他的恶劣的工作感到羞愧；它只是某个不独立的，低级的神的作品；对于较高级的神，它是嘲笑的对象；它是某个老迈的神的衰朽期的作品；自从他死了之后，它就依靠着它从神取得的起始的冲动和动力，往前乱撞乱碰。[40]

休谟设想了一幅多神论的情境，这些神的资质良莠不齐，创造出来的宇宙质量高下不等，就像徒弟试图模仿师傅的作品。但如果我们将“幼稚的神”和“老迈的神”代之以各种模拟者的话，那么他所预想的就是一个遍布虚拟宇宙的世界，其中有好的、前景光明的，也有次的、难逃厄运的。

如果所有可能的世界都存在，而我们活在一个自然法则与众不同的虚拟世界中，这和只有一个真实的世界有什么区别吗？实际上，这应该有什么区别吗？[41]如果你是一个想要研究世界运转规律的（虚拟）科学家，那么这会让你感觉相当失落，毕竟任何事情都可能无缘无故地发生。所以毫不意外，在科学世界观里，虚拟现实并不受欢迎，因为它们的存在会动摇科学世界观的根基。哲学家们倒是更能严肃地考虑这样的问题，有些甚至会以此为背景讨论伦理学问题。他们提出的一些问题发人深省。

罗宾·汉森指出，如果存在虚拟现实，那它就会以一种独特的方式影响你的行为。[42]虚拟的经历，无论看起来多么真实，都比真实的经历更可能遭遇无法预料的戛然而止。这让汉森推断出，“在其他条件相同的情况下，你会更不在意别人，更愿意活在当下”。我们很熟悉电影和舞台剧中这样的场景，明星的周围簇拥着大批优秀演员，他们和明星之间有互动；再往远处，就是临时演员和打零工的低薪演员，他们以低成本充当着背景。同理，在虚拟现实之中，离你很远的人物可能不过是一些虚假角色，你不必太在意他们。汉森指出，如果你处于他人的模拟之中，那么最重要的是，你要让自己更富有娱乐精神！要出名！要变成重要角色！这会增加你在虚拟现实中的存活几率以及将来别人再次把你模拟出来的机会。而如果你没能拥有这些特质，就会像肥皂剧中的角色一样，被编剧发往符拉迪沃斯托克度长假，然后再也不会回来了。

如果我们看看新闻里那些人的行为举止，再想想汉森的话，也许就很容易得出，我们一定生活在模拟之中。然而，以上这些都不太有说服力。你应该如何举动完全依赖于模拟者的道德立场。如果他们喜欢娱乐，你就会变得富有娱乐精神。而如果他们追求一种崇高的目的，那么通过为正义事业牺牲殉道，你就非常有可能被再次模拟出来。虽然我们并不认为这些结论应该成为你生活中的行为规范，但它们确实凸显了道德哲学及其实践的核心问题。如果虚拟现实很常见，而我们又处于其中之一，而且它又是我们熟知的那种模拟的话，麻烦就大了。但是凭什么虚拟现实就得是我们熟知的那种模拟呢？如果我们总是用“模拟”这个词来描述上帝创造一次性宇宙的行为结果，那么我们就生活在一场非常类似的模拟之中，尽管模拟者属于另一种类型。

也有些人认为，从我们可能生活在虚拟现实中的想法得出的这些异乎寻常的推论，反而成为了否定其他世界存在的强有力论据。如果绝大多数世界是虚拟的，那么这些世界之中就会存在虚假的物理定律，而我们就会犯下另一种滑坡谬误：因为不存在任何可靠的知识，所以我们什么都不知道。这是与唯我论①恰好相对的一种论调，也会导致类似的、使思想瘫痪的后果。如果可能性有无限多种且都是存在的，那么现实无疑远远超出了我们的承受能力。

① 唯我论是指，除了我和我的思想之外，其他一切事物都不可信。——译者注

没有新鲜事的宇宙

……图书管理员推断出，这个图书馆“无所不包”，书架上的书包含了25个符号［22个字母、空格、逗号、句号］的所有可能的组合。是的，所有一切——巨细无遗的未来历史，第八级天使的自传，该图书馆的正确书目，成千上万的错误书目，证明了这些书目的错误之处的手

稿，证明了正确书目的错误之处的手稿，巴西里德斯（Basilides）的诺斯替教派福音书，该福音书的注解，该福音书注解的注解，关于你的死亡的真实记录，所有书的所有语言译本，所有书的所有改写版本，比德（Bede，672～735）原本可能写（但事实上没写）的撒克逊人的神话，塔西佗的那些失传了的书。

——豪尔赫·路易斯·博尔赫斯（1899～1986，阿根廷作家）[43]

试想，生活在一个没有新鲜事的宇宙。一切都是赝品。没有一个想法是新的。不存在新鲜感，也不存在创新性。没有什么东西是第一次出现的，也没有什么东西是最后一次出现的。没有什么是独一无二的。每个人都有自己的副本，而且不是一个副本，是无穷无尽个副本。

如果宇宙的空间无限宽广且存在生命的概率大于零，那么这种奇闻异事就有可能发生。这是因为无限大与任何有限的大数目存在本质上的不同，而不论那个大数目有多大。[44]

在一个体积无穷大、物质无穷多的宇宙中，任何发生概率不为零的事情一定会发生无穷多次。[45]这样此时此刻，我们每个人都会有无数个一模一样的副本，他们完全模仿着我们的一举一动。每个人也有另外一些一模一样的副本，他们在忙一些别的事情。事实上，每个人都有无数个副本在忙活任何我们此时可能忙活的事情，只要其概率不为零。这种惊人的事态被称为“副本悖论”（replication paradox），德国哲学家弗里德里希·尼采（1844～1900）在《快乐的知识》一书中讨论了这个问题，当时人们发现了无限大的宇宙会导致这样的后果。尼采写道：

在这个巨大的几率（这是宇宙的存在基础）游戏中，宇宙会经历可计数的种种组合状态……如果宇宙无穷无尽，那么总有一个时刻，所有可能的组合都出现了，而且出现不止一次，而是无穷多次。[46]

这个副本悖论会带来种种古怪的推论。我们相信生命演化的概率大于零（因为我们存在），因此在一个无限大的宇宙中，必然存在无穷多个存活的文明。在这些文明之中，存在着我们不同年龄的副本。一处的我们死去了，还有别处的无数个我们的副本，他们都有跟我们相同的记忆、相同的经历，并且会继续活下去。这种传承会无限延续下去，所以从某种意义上说，我们都在这个悖论情境中得到了“永生”。这进一步引发了另一个古怪的推论。如果考虑我所有可能的经历，其中的我都有相同的过去，未来的命运却充满各种可能，那么在下一瞬间，死去的我都会远远多过活着的我。于是，为什么我依然活着呢？

这个悖论也进入神学讨论，并引发了巨大争议。假设我们将同样的推理过程应用到“道成肉身”的问题中。在一个无限宽广的宇宙中，如果一件事的发生概率有限大，那么这件事必然已经在别的地方发生过无数次了。奥古斯丁（354～430）[47]在4世纪时就用这种逻辑声称，地球上的有感情的生命一定是独一无二的，否则耶稣殉难于十字架的事件就要在其他世界中发生无数次了。[48]而 18 世纪下半叶时，托马斯·潘恩（1737～1809）则主张，其他地方也存在生命的说法明显是对的，因此，十字

架殉难事件并没有发生（或者至少没有它所声称的影响），因为让它发生无数次实在太荒谬了。

我们可以问，如果我们见到了自己的一个副本，会有什么后果呢？你也许会想，这就像对着镜子进行空拳练习一样，但事实上没有理由认为，你的任何一个副本都会复制你的全部行为。在相遇之前，你们可能拥有完全相同的经历，但当你们面对这个新的处境时，你们可能会做出不同的反应，就像两个双胞胎那样。这是你们第一次分道扬镳，而在未来，你们的经历和选择会越来越不同。然而在无限宇宙的其他地方，存在着我们的无穷多个副本，他们与我们一模一样，也要面对同样的决策问题。这就好像我们在每一时刻所有可能作出的选择都被实现了。总有一个地方存在这样一个人，他的过去经历和我完全相同，并会不断从我可能作出的选择中选出一个，进行下一步行动。作出了跟我不一样选择的人总是比跟我作出相同选择的人多得多。

这种“理论”的最有意思的地方在于，如果它是正确的话，那么它不可能是原创的。从前，它已经被提出无数多次了。事实上，如果一个无限的宇宙尝试了所有的可能性，那么没有一样东西会是新鲜的。

一些宇宙学家认为这样的景象太过骇人，因而把这作为有限宇宙的有力论据。[49]还有一些宇宙学家认为，如果一个宇宙允许所有可能的事件发生而不计较其后果，那么由此得出的道德意涵令人堪忧。一切可能的或然历史都变成了现实。在这些历史之中，有时候邪恶总是能够战胜正义，而有时候邪恶甚至被当成了正义。

在现实中，物理定律为我们提供了某种保护，使我们免遭其中一些恼人的后果。即使宇宙总体上是无限的，我们也只能与其中有限的部分发生联系，因为光速是有限的。今天，我们所能接触到的最远距离大约是 10^{27}米。相比之下，地球的大小是1.3×10^{7}米，到最近的恒星的距离约为6×10^{16}米。它们都非常遥远，但比起见到自己的副本所需航行的距离，就是小巫见大巫了。为了有机会见上自己的副本一面，你需要走上 $^{1028}10$ 米；而为了见到地球和其中一切事物的副本，你需要走上 $^{1030}10$ 米。如果走上 $^{10120}10$ 米远，你就有50%的几率见到可见宇宙的所有副本当中离我们最近的那一个。光速的有限性使得我们无法在无限宇宙中见到自己的副本。但即使我们和副本之间相距甚远，仍然会有一些关于副本的问题令人深感不安。

如果无限的宇宙总是大致处于相似的状态，就像旧的稳态宇宙模型一样，那么就会出现有关时间的悖论。任何事情，如果发生的概率是一个有限的值，就会在过去的历史中发生无穷多次。这个宇宙之中不存在任何新鲜感。此外，因为演化出智慧生物的概率是有限的，所以这样的生物必然是无限普遍的，随着时间的推移，这些生物会在宇宙中大量繁

殖。我们应该对这样的生物司空见惯了才对。这个结论如此古怪，再加上缺乏地外文明存在的证据，于是保罗·戴维斯[50]、弗兰克·梯普勒和我[51]以此作为反对稳态宇宙和静态宇宙支持者的观测证据。当然，所有关于无限宇宙的结论，如果正确的话，都不可能是原创的。

最后我还想提及一个严肃而发人深省的想法。我们对无限复制的想法心存疑虑，因为它看起来既奇妙，又荒谬，而且似乎不可能出现。但其实我们的周围就存在着这样一类完全相同的复制品，世界就是由它们组成的。质子、电子、夸克，所有这些自然界的基本粒子被称为全同粒子。一旦你看到一个电子，就已经看到了所有的电子。[52]没有人知道为什么会这样。宇宙建立在复制的基础上，我们怀疑这种复制品的数量是无限的，就像膨胀的宇宙看起来是无限的一样。这是所有微调过程中最美妙的一个。可惜大多数物理学家甚至没有注意到这一点，也没有对此留下什么评论。但这意味着，复制是构成现实的深层结构的核心。

玻尔兹曼的宇宙

我相信，我们正以不可逆转的趋势奔向自由和民主——但情况也可能变化。

——丹·奎尔（Dan Quayle，老布什任期的美国副总统）

无限宇宙种种令人不安的意涵，在极端条件下考验着我们的科学方法以及我们对其的信心。另一个计算多重宇宙中存在观测者的可能性的方法是，计算某个观测者在宇宙中观测到某些结果的概率。可惜，仅仅这样还不够。如果无限宇宙中产生观测者的概率是有限的，那么他们生活的地方会有无穷多个，但在不同的情况下，他们看到的东西会以不同的概率出现。由于存在无穷多个观测者的副本，我们就无法准确地预测出，最有可能出现的是哪些观测结果。

说到“观测者”，通常会有一种批评说，我们局限于用一种太过沙文主义的思维方式来定义生命。我们只能把地外生命想象成有点儿类似我们的样子，由原子组成，有大脑，有肉体；他们也许像计算机，但这仍然是地球上的信息处理装置的一种外推。这种定义生命的考量与 SETI（地外文明搜索计划）[53]的搜索策略紧密相关。

从前，人们往往通过探测无线电信号来搜索地外文明的活动，或者高科技活动的迹象。所谓高科技活动的迹象，意味着高耗能、高混乱度、改变行星的轨道或者影响恒星的演化，所以可能被我们看到。1964年，苏联天文学家尼古拉·卡达谢夫（Nikolai Kardashev）甚至用文明所利用的能量形式来划分文明的等级，如是否能操纵行星、恒星以致整个恒星系。[54]我增加了Ⅳ型文明，也就是能够操纵整个宇宙，就像我们在本章前面看到的那样。[55]然而，高级文明更可能的是已经将它们的技术微型化了。我们自己的技术就是在稳步地向微型化迈进。量子技术的研究使我们得以操纵单个原子，并用少量的原子和分子制造机器。纳米技术和量子计算机技术激动人心的发展使得在微小的空间内存储和运行海

量信息成为可能。考虑到高科技生物会面临环境退化和资源枯竭的问题，我们可以猜测这样的高级文明会不断推动技术的升级换代，进一步微型化。文明的等级水平或许更应该由操纵微小尺度的能力来衡量，而不是操纵巨大尺度的能力。[56]这意味着文明越高级，他们的技术活动就越难以察觉，因为他们的能耗更低，产生的废热更少。他们的空间探测器甚至会比一团原子或一团分子还要小。我们根本无法注意到这些探测器的存在。

宇宙学家已经考虑到了更加离奇的“文明”形式。正如我们在第 2章提到过的，1895年，早在膨胀宇宙的理论远未提出之前，奥地利物理学家路德维希·玻尔兹曼就已经开始冥思苦想，为什么我们发现无序度在随着时间的推移而增加（我们称之为热力学第二定律，或者“时间之箭”），而牛顿的运动定律却总是允许观测到的事件以相反的顺序逆着时间线发生。[57]

现在，我们已经知道，物理学定律既允许酒杯从桌子上跌落，碎了一地的解，也允许这个解的时间反演，也就是一地碎片自动组合成完整的酒杯。但后一种解所需的初始条件极难满足，所以我们从未见过这样的事情发生。无序转化为有序是一种时间反演的过程，要求必须满足极为苛刻的初始条件，这样所有的碎片都大小适当，运动方式也恰到好处，恰好同时拼凑成一个完整的杯子。这就像是把一万块拼图板扔出去，发现它们落地时刚好组合成了一幅完整的图案一样。

玻尔兹曼想知道，宇宙的不同部分是否出现或演化出了热力学中不同程度的不可能的状态。温度处处相等的状态是最可能的，而任何自发地偏离这种均匀性的状态就不那么可能出现了。他设想，宇宙大体上平均而言处于热平衡状态。但为了使生命能够出现，宇宙的不同时期必然存在一些偏离了平衡状态的涨落。在其中一些涨落之中，有序度会随着时间的推移而增加，而在另一些涨落中有序度则会减少，两种涨落的多寡体现出两种初始状态所对应的不可能程度。玻尔兹曼和他的学生恩斯特·策梅洛（后来成为著名的理论数学家）认为，我们必然生活在其中一个寿命相当长的涨落之中，随着时间的推移，它的有序度在减少。

关于无序度随着时间增加的原因，玻尔兹曼提出了两种可能的解释。第一个是说，一开始世界就处于高度有序的状态，因此，之后最可能发生的事情就是走向无序（我们看到酒杯摔成碎片）。另一个是说，一开始宇宙的不同区域就有不同的有序度和无序度，分别表现出不同的行为；有一些更有序，有一些更混乱，但是生命必须存在于一个从有序出发走向无序的区域中。玻尔兹曼设想了无穷多个独立的世界，每个世界都形成于平衡状态的一次罕见的涨落，世界和世界之间相隔千山万水——要比从太阳到最近的恒星，如天狼星的距离还要远 10^{100}倍。他没有

告诉大家他更喜欢哪一种解释，甚至还将第二种解释归功于他多年的实验助手（舒茨博士）。[58]我们在第2章中说过，实际上这种解释十多年前就由英国工程师、物理学家塞缪尔·托尔弗·普雷斯顿在各种场合中提出了。[59]普雷斯顿后来到德国攻读博士学位，师从玻尔兹曼，并于1894年获得了学位。我很想知道这个神秘的“舒茨博士”是否真的就是塞缪尔·托尔弗·普雷斯顿。

第二个人择涨落的解释吸引了一些有趣的评论者。著名的法国数学家、物理学家亨利·庞加莱（Henri Poincaré，1854～1912）就喜欢这个解释，因为如果人类生存在一个非典型的涨落当中因而最终的大混乱并非不可避免的话，那么这就可能使人类免遭宇宙的熵达到最大时的终极热寂。英国生物学家霍尔丹（J. B. S. Haldane，1892～1964）则认为，适合生命存在的涨落的出现几率极低，因而地球上的生命应该是独一无二的。因为如果附近还存在相似的生命形式的话，我们就不得不解释为什么会出现这么大（更大就意味着更不可能出现）、大到超出让地球产生生命所需的一个涨落。换句话说，霍尔丹关心的是，其实我们只需一个无序度在逐渐增加、能够维持生命进而产生观测者的涨落。这个涨落只要跟太阳系一样大就够了。这比宇宙的尺度小得多，因而更可能出现，而我们已经观测到宇宙在大尺度上是遵守热力学第二定律的。[60]在人择涨落的解释里，想要让整个可见宇宙无序度都像我们在地球所见的一样不断增加是极为不可能的。

这个古老的热力学问题在现代宇宙学中留下了两个残留的影响。一个是罗杰·彭罗斯提出的，宇宙现在正处于一种极端有序的状态中，所以140亿年以前的宇宙应该处于一种更加苛刻的有序状态之中。[61]宇宙学家不太相信这个说法[62]，因为测量宇宙无序度（或“熵”）的大小就等于测量了宇宙的年龄，任何产生了恒星并形成了碳元素的古老宇宙必然伴有一个巨大的熵。更重要的是，暴胀就是一种产生局部有序度的方法，它能将无序赶出我们的视界之外，这样它们就看不到了。我们并不知道宇宙在整体上是否处于低熵有序的状态。我们只知道可见宇宙的这一部分如何，因此不能推断出整个宇宙的初始状态都是高度有序的，或者说现在整个宇宙都处于一种极不可能出现的状态之中。

最近，玻尔兹曼的解释又在宇宙学研究中再次出现。他的人择涨落意味着，为了能够产生文明和像我们这样的智慧生物，偏离平衡状态的涨落就得足够大，而这种几率是极低的，因而宇宙中一定充满了更小、有序度更低、无躯骸且转瞬即逝的智慧生物。在这个解释中，这样的生物比我们更可能出现，并且在一个无限的宇宙中应该已经随机出现了无穷多次了。

这些生物被称为“玻尔兹曼式大脑”。由此我们要面对这样一个悖

论：在宇宙中，随机产生玻尔兹曼式大脑及其黄粱一梦的虚假记忆的几率，比起产生数十亿像我们一样拥有自我意识以及错综复杂真实记忆的生命的几率要大得多。[63]然而，正如我们在这一章所说的，尽管宇宙可能是无限的，并在某种意义上处于无限的随机过程之中，但这并不是事情的全部。存在物理规律主宰着事物变化的方式，还存在物理机制驱动了宇宙的快速膨胀，使得我们只能看到无限多重宇宙的某一个小角落，并且可能是个有点不一般的角落。这些进展，不论它们是多么纯理论，它们都清楚地揭示了一点，即我们的存在以及其他智慧生物的存在，都是我们在思考宇宙和检验理论时所必须考虑的因素。观测者对宇宙学理论很重要，对宇宙也很重要。

注释

[1] 最大的难点在于，最终的答案必须与得出答案的观测者的运动状态无关。爱因斯坦的理论认为，同时性并不是一个与观测者无关的绝对概念。在我看来是同时发生的事情，在你看来就可能不是同时发生的。于是，计算各个地方同时产生不同事物的概率也面临着这种模棱两可的局面。如果想绕开这个问题，直接算出一个泡泡宇宙可能经历的所有历史又会面临另外一种难题：最终的答案依赖于这个宇宙最初设定的真空能。人们又发现了一种可靠的方法，也就是将多重宇宙的数学性质投影到一个高维的“屏”上，每个将来会膨胀成宇宙的泡泡在屏上所占的大小很容易计算。令人惊讶的是，这两种方法之间似乎存在某种深层的联系，而且它们都能相互弥补对方的不足。

[2] 计算宇宙或多重宇宙中任何物理机制不同结果所具有的概率分布的问题，都叫做“测度问题”（Measure Problem）。一般而言，我们现在还不知道如何计算出宇宙每一种基本性质的概率分布。

[3] M. Druon,*The Memoirs of Zeus,* Charles Scribner’s and Sons, New York (1964).

[4] 宇宙的膨胀持续了多长时间取决于暴胀持续了多长时间。

[5] P. A. M. Dirac, ‘Reply to R. H. Dicke’,*Nature* 192, 441 (1961). 1980年，我收到了狄拉克为这个问题亲笔撰写的说明，其中的措辞和他1938年首次提出时的一模一样。

[6] 转引自：G, Farmelo,*The Strangest Man*, Faber & Faber, London (2009), p.221.

[7] J. D. Barrow, ‘Life, the Universe, but not quite Everything’,*Physics World*, Dec., pp. 31–35(1999).

[8] M. J. Rees,*Comments on Astrophysics and Space Physics.* 4, 182 (1972); M. Livio,*Astrophy. J.* 511, 429 (1999).

[9] 关于这个理论的具体内容和相关研究在1986年之前的发展历

史，参见：J. D. Barrow and F. J. Tipler,*The Anthropic Cosmological Principle, Oxford University Press*, Oxford (1986).最近有关变动自然常数的内容，参见：J. D. Barrow, *The Constants of Nature,* Jonathan Cape, London (2002).

[10]英国神学家弗里德里希·泰南特（F. R. Tennant，1866～1957）在他1930年出版的《哲学神学》第二卷（*Philosophical Theology,* vol. 2, Cambridge University Press, Cambridge, p. 79）中首次将“人类的”（anthropic）一词引入了宇宙学。泰南特设想了能够作用在整个宇宙尺度上的设计和各种类型的目的论，借此这个宇宙将落入“人类的范畴”（anthropic categories），从众多可能的世界中脱颖而出，得以允许智慧生物存在。关于泰南特的成果，更详细的描述参见：J. D. Barrow and F. J. Tipler, *The Anthropic Cosmological Principle,* Oxford University Press, Oxford (1986), section 3.9.

[11] B. Carter, ‘Large Number Coincidences and the Anthropic Principle in Cosmology’, in M.S. Longair (ed.), *Confrontation of Cosmological Theories with Observational Data*, IAU Symposium, Reidel, Dordrecht (1974), p. 132.卡特第一次在现代天文学中提到了“人择原理”（anthropic principle）一词，不过以前也有一些哲学家提到过这个词。

[12]在杆状结构可以拼接成的二维凸多边形中，只有三角形是刚性的。这解释了为什么高压线铁塔是由相互嵌套的三角形结构组成的，为什么有两个横栅的门很常见。更多内容，参见：J. D. Barrow, *100 Essential Things You Didn’t Know You Didn’t Know, Bodley* Head, London (2009), chapter 1.

[13]关于这封信的更多内容，参见：J. D. Barrow, *The Constants of Nature,* Jonathan Cape, London (2002), chapter 3.

[14] H. Mankell,*Chronicler of the Winds,* Harvill & Secker, London (2006), p. 25.

[15] B. Russel ,*Logic and Knowledge,* ed. R. C. Marsh, Alen and Unwin, London (1956).

[16] 在伏尔泰的小说《老实人》中，他刻画的潘格洛斯（Pangloss）博士正是讽刺了这种莱布尼茨式的乐观主义——“我们生活的世界就是所有可能世界中最好的一个”。一个显而易见的批评是，我们既不知道“其他世界”是什么样子的，也不知道所谓“最好”是什么意思。对此法国数学家皮埃尔·莫佩蒂（Pierre Maupertuis，1698～1759）在1746年提出了辩驳。不满于自然神学家们常常借助这些模糊的概念，用以证明上帝已经为我们准备好了一切，他给出了一套明确的说法，来

解释什么是最好的世界。莫佩蒂首次证明了，有两种方法能够确定物体的运动路径。一种方法是，可以借助牛顿的运动定律，确定物体的初始位置和速度，然后通过求解方程来确定物体未来的运动路径。另一种方法是，假设物体可以沿着初始时刻初始位置和最终时刻最终位置之间的所有可能的路径运动。选择实际运动路径的方法是，让某一个包含了质量、速度和距离的物理量，也就是“作用量”，取到极小值。这样做了以后，人们总能得到和牛顿定律相同的结果。实际上，从“最小作用量原理”出发可以推导得出牛顿的运动定律（爱因斯坦的理论也是如此）。莫佩蒂认为，作用量没有达到极小值的路径就是莱布尼茨的批评者所梦寐以求的“其他世界”，而所谓“最好”正好说明自然界中实际路径的作用量达到了极小值。19世纪时，一些法国时事评论员竟然想把新发现的化石与这些非最小作用量的、被淘汰了的世界联系起来，他们认为其中的生命已经灭绝了。莫佩蒂在 1750年的书《论宇宙》（*Essai de cosmologie*）中的深入论述，可参见：Barrow and Tipler,*The Anthropic Cosmological Principle*, section 3.4.关于可能世界的现代哲学观点，可参见戴维·刘易斯关于模态实在论（Modal realism）的研究。刘易斯认为，所有可能世界像真实世界一样都是真实的，因为所有可能世界和真实世界是同一类事物，不可能进一步化简为更基本的元素；但这些世界之间是相互隔离的，所以不存在因果联系。可参见：D. Lewis,*Convention: A Philosophical Study*, Harvard University Press, Cambridge, Mass. (1969); D. Lewis, Counterfactuals, Harvard University Press, Cambridge, Mass. (1973, rev. edn 1986).

[17] S. K. Blau and A. H. Guth in *300 Years of Gravitation*, eds. S. W. Hawking and W. Israel, Cambridge University Press, Cambridge (1987).

[18] E. R. Harrison, ‘The Natural Selection of Universes Containing Intelligent Life’,*Quart. J. Roy. Astron. Soc.* 36, 193 (1995).

[19]哈里森给论文起的标题叫“宇宙的自然选择”，但实际上他提出的机制跟宇宙的强行孵化非常类似。

[20] D. Adams, *The Original Hitchhiker Radio Scripts*, ed. G. Perkins, Pan Books, London(1985).引自其中的《第七集》（‘Fit the Seventh’），于 1979年 12月 24日在BBC第四套广播中首播。

[21] L. Smolin,*Class. Quantum Gravity* 9, 173 (1992); L. Smolin,*The Life of the Cosmos*, Oxford University Press, Oxford (1996).

[22]可参见：J. A. Wheeler, ‘From Relativity to Mutability’, in J. Mehra (ed.), *The Physicist’s Conception of Nature*, Reidel, Boston (1973), pp. 239 ff; C. Misner, K. Thorne and J. A. Wheeler,*Gravitation*, W. H. Freeman, San Francisco (1973), p. 1214.

[23]注意，这里说的繁殖是一种理论上的可能，而不像永恒暴胀宇宙那样，是指实际的物理过程。

[24]一个更具体的例子，贝肯斯坦和霍金发现，黑洞的引力熵正比于GM^2。所以如果我们忽略G的变化，当黑洞的质量M增加时，得到的引力熵也会随之增加，这就符合斯莫林的假设。但如果G也会变化的话，即使M减小，如果G的增加速度非常快，也会使引力熵随之增加。因此，我们不应该假设，宇宙在经历过许多次坍缩和重选参数的过程之后一定会导致黑洞产生率的最大化。

[25] R. Hanson, ‘How to Live in a Simulation’, *Journal of Evolution and Technology* 7 (2001),http://www.transhumanist.com.

[26] J. D. Barrow,*Pi in the Sky: Counting, Thinking and Being,* Oxford University Press, Oxford (1992), chapter 6.

[27] N. Bostrom, ‘Are You Living in a Computer Simulation?’, http:/www.simulation-argument.com.

[28]但是这又回到了我们在本章前面强调过的尴尬的概率测度问题。博斯特罗姆的论述包含了一个隐含假设，即无论什么类型的世界，真实的也好，虚拟的也好，出现的概率都大致相等，至少差别不会太大。但实际上不是这么回事。

[29] P. C. W. Davies, ‘A Brief History of the Multiverse’,The New York Times, 12 April 2003.

[30] L. Susskind, *The Cosmic Landscape: String Theory and the Illusion of Cosmic Design,* Little Brown, New York (2005); A. Vilenkin, *Many Worlds in One: The Search for Other Universes,* Hill and Wang, New York (2006).

[31] J. K. Webb, M. Murphy, V. Flambaum, V. Dzuba, J. D. Barrow, C. Churchill, J. Prochaska and A. Wolfe, ‘Further Evidence for Cosmological Evolution of the Fine Structure Constant’, *Phys. Rev. Lett.* 87, 091301 (2001).

[32] S. Wolfram,*A New Kind of Science,* Wolfram Inc., Champaign, Ill. (2002).

[33] K. Popper,*Brit. J. Phil. Sci.* 1, 117 and 173 (1950).

[34] D. MacKay,*The Clockwork Image*, IVP, London (1974), p.110.

[35] J. D. Barrow, *Impossibility,* Oxford University Press, Oxford (1998), chapter 8.

[36]不过，赫伯特·西蒙有一篇著名的高引用率的文章提出了相反的观点，但其论证是错误的，参见：Herbert Simon, ‘Bandwagon and Underdog Effects in Election Predictions’, *Public Opinion Quarterly* 18, Fall,

245 (1954); reprinted in S. Brams, *Paradoxes in Politics,* Free Press, New York (1976), pp. 70–77.文中的错误在于，作者不恰当地在离散变量，而不是连续变量中应用了布劳威尔不动点定理。更详细的说明参见：K. Aubert, ‘Spurious Mathematical Modelling’, *The Mathematical Intelligencer* 6, 59 (1984).

[37] J. D. Barrow,*The Constants of Nature: From Alpha to Omega,* Jonathan Cape, London (2002).

[38] 1965年，英特尔公司的创始人之一戈登·摩尔发现，每过两年，每平方英寸集成电路上的晶体管都会进一步微型化，数量翻倍而成本减半。这个论断在一定精度下可以预测技术的发展趋势。其他文明的技术微型化趋势也可能存在某种类似的规律。摩尔定律对计算机产业的发展非常重要，因为软件公司和硬件公司可以根据这个规律调整自身的发展战略，以跟上技术发展的脚步。

[39]关于这些宗教论证以及休谟的回应（及其对康德的影响）的深入讨论参见：Barrow and Tipler,*The Anthropic Cosmological Principle,* chapter 2.

[40] David Hume, *Dialogues Concerning Natural Religion* (1779), in Thomas Hill Green and Thomas Hodge Grose (eds.),*David Hume: The Philosophical Works,* vol. 2, London, 1886, pp. 412–416.

[41]这些问题与瑞·库兹韦尔（Ray Kurzweil）在《灵魂机器的时代》一书［*The Age of Spiritual Machines,* Viking, New York (1999)］中讨论的主题密切相关。库兹韦尔思考了虚拟现实和人工智能的产物拥有心灵和审美能力的可能性。

[42] Hanson, ‘How to Live in a Simulation’.

[43] J. L. Borges,*The Library of Babel,* D. Godine, Jaffrey, NH (2000; 1941年以西班牙语首次出版).

[44] J. D. Barrow, *The Infinite Book: A Short Guide to the Boundless, Timeless and Endless,* Jonathan Cape, London (2005).

[45]这种无穷大宇宙的各方面性质必须是彻彻底底随机产生的，这样才会发生副本“悖论”。如果宇宙的体积无穷大却只包含一个原子的话，情况就不满足了。

[46] F. Nietzsche, *The Gay Science: With a Prelude in Rhymes and an Appendix of Songs* (first publ. 1882); translated, with commentary, by Kaufmann, Vintage Books, London (1974).

[47]无限重现、无限循环的思想出现得远比奥古斯丁早得多。大约在公元前350年，罗德岛的欧多莫斯（Eudemus of Rhodes）将这个想法归功于毕达哥拉斯学派：“假如有人赞同毕达哥拉斯学派的观点，即一

模一样的事情还会再次发生，那么我还会手执教鞭、像你一样端坐，再跟你讨论一次，而且其他每一件事情都跟现在一模一样。”参见：G. S. Kirk and J. E. Raven, *The Pre-Socratic Philosophers*, Cambridge University Press, New York (1957), Eudemus Frag. 272.

[48]关于这个矛盾有很多富有想象力的回答，例如 C. S.刘易斯（1898～1963）的科幻三部曲《沉寂的星球》（1938）、《皮尔兰德拉星》（1943）和《黑暗之劫》（1945）中，他假设地球是宇宙中道德最败坏的地方，而其他有生命的世界并不需要上帝来救赎。

[49] G. Ellis and G. B. Brundrit,*Quart. Journal. Roy. A stron. Soc.* 20, 37–41 (1979).

[50] P. C. W. Davies,*Nature* 273, 336 (1978).

[51] Barrow and Tipler,*The Anthropic Cosmological Principle*, section 9.5.

[52]当理查德·费曼向约翰·惠勒指出这个现象时，惠勒回应说，也许因为宇宙中只有一个电子。

[53]最近的研究参见：P. C. W. Davies,*The Eerie Silence: Are We Alone in the Universe*, Allen Lane, London (2010).

[54]对Ⅰ型文明来说（比如我们的文明），它能够驾驭一个行星的能量，即10^{17}瓦特的功率。Ⅱ型文明可以驾驭一颗恒星的能量，即10^{26}瓦特。Ⅲ型文明可以驾驭一个星系的能量，即10^{37}瓦特。因此，一个文明的类型可以粗略地用公式“类型=($\log_{10}$P-6)/10”来估算，其中P是文明所能驾驭的能力。这个公式是卡尔·萨根首先提出的，参见：*Cosmic Connection: An Extraterrestrial Perspective*, ed. J. Agel, Cambridge University Press, Cambridge (1973).

[55] J. D. Barrow, *Impossibility*, Oxford University Press, Oxford (1998), pp. 129–131.

[56]我将之标记为-Ⅰ型、-Ⅱ型，依此类推。参见：Barrow, *Impossibility*, pp. 133–138.

[57] Barrow and Tipler,*The Anthropic Cosmological Principle*, section 3.8.

[58] L. Boltzmann,*Nature* 51, 413 (1895).

[59] S. T. Preston,*Nature* 19, 462 (1879), and *Philosophical Mag.* 10(5), 338 (1880).

[60]对此的另一种表述参见：R. Feynman, *The Character of Physical Law*, MIT Press, Cambridge, Mass. (1965).

[61] R. Penrose,*The Road to Reality*, Jonathan Cape, London (2004).

[62]彭罗斯的说法利用了黑洞的引力场伴有巨大的量子熵的特点。

可以想象，如果物质都以巨型黑洞的形式存在，宇宙所包含的熵就会多得多。最大的黑洞所伴有的熵也是最大的，因为黑洞熵正比于黑洞的表面积。最大黑洞的大小不可能超过宇宙的年龄 t乘以光速，所以最大黑洞的表面积和熵正比于t^2。

[63] L. Dyson, M. Kleban and L. Susskind,*J. High Energy Phys.* 0210, 011 (2002); A. Linde,J. of *Cosmology and Astroparticle Phys.* 0701, 022 (2007); D. N. Page, *J. Korean Phys. Soc.* 49, 711 (2006).

第11章 非主流的宇宙

理论啊理论，多得数不胜数，像一阵秋风经过，落木萧萧而下，像暴风掀开了巨大的造纸厂，碎屑纷纷而上，又像黑压压的尘云，挟裹在思想的飓风里。在这场席卷一切的干旱中喘着粗气，我差点忘了，每一粒尘埃之中都会带有一些真理的种子，大部分最后都干枯死亡了，但也有一些会活下去，产生更多的种子，变得至关重要。

——奥拉夫·斯德普尔顿（1886～1950，英国科幻小说家）[1]

黏合拓扑的宇宙

雷丁是个很有意思的地方，甚至连最推崇雷丁的人都喜欢提到它离别的地方很近。“从伦敦市中心坐火车到这里只要25分钟”，值班经理热情洋溢地说。“做大巴去希斯罗机场只要45分钟！”购物中心的咖啡店女服务员说。

——英国航空《商业生活》杂志[2]

如果宇宙能从“无”中产生而不违反物理定律的话，也许用这些定律就可以描述宇宙的创生了。相信宇宙存在开端的宇宙学家都习惯认为，物理定律在那个时刻失效了。物理定律是和空间、时间和实体的宇宙一起冒出来的。我们所有关于“其他”宇宙和多重宇宙的讨论则对这一观点提出了质疑。渐渐地，我们发现我们的宇宙只不过是自然法则的一个产物，也许只是多重宇宙的一部分：永恒历史中的一小段局部事件。尽管这种想法有点贬低了环绕我们四周的广袤宇宙，但它同时又抬高了自然法则的地位。自然法则似乎容许更多宇宙的存在。整个20世纪中，宇宙学家都在研究从爱因斯坦方程组中得到的各式各样的宇宙模型，可是他们又希望会有某种特别的原理，或是初始状态，可以让他们从中挑出一个最真实的宇宙模型来。现在，出人意料地，我们发现在多重宇宙中可以存在很多可能的宇宙，甚至是所有可能的宇宙。

从无中最有可能产生哪种类型的宇宙呢？所有遵守爱因斯坦方程组的宇宙的能量和电荷似乎都是零。这两个物理量和另外一个物理量在所有物理过程中都保持守恒，是最重要的三个物理量。能量和电荷都可以到处流动、重新分布，但如果你分别计算它们的总和，把正值和负值加在一起，最后的结果一定是不变的。物体之间的差异还体现在第三个重要的守恒量，也就是角动量上。这个量度量了物体的旋转程度，而我们已经知道爱因斯坦方程组虽然要求能量和电荷是零，但并没有要求角动量也是零。哥德尔发现可能存在旋转的宇宙，但是我们还没有找到任何与宇宙总的旋转程度有关的证据。而且我们也没有指望能找到这样的证据，因为如果宇宙发生过暴胀，那么暴胀就会将任何初始的旋转减弱到无法观测的水平。[3]

1973年，美国粒子物理学家爱德华·特赖恩[4]设想了这样一种情况（帕斯库尔·约尔丹和乔治·伽莫夫以前也考虑过[5]）。他提出，宇宙也许不过是量子真空中的一个虚拟涨落。量子真空会受到海森堡不确定性原

理的限制。这个原理认为，我们无法同时知道自然界中任何事件的位置和动量的精确信息，也无法同时知道时间和能量的精确信息。在这个图景中，特赖恩认为我们的宇宙可能“只不过是间或发生的那些事中的一员”。不确定性原理关于能量–时间的表述是，一个涨落持续的时间和能量的乘积大于一个物理常数（$\Delta t \times \Delta E > h$）。[6]“这说明一个能量为零的涨落——比如我们的宇宙——可以拥有无穷长的寿命。”涨落最有可能的寿命实际上是 10^{-43}秒。微小的、转瞬即逝的涨落如果要变得像宇宙一样广袤，一样持久，那么就必须有暴胀作用在涨落之上。特赖恩没能解释为什么我们的宇宙存在了如此之久，因为八年之后才会出现暴胀理论。他所能做的就是利用长寿的大型真空涨落的稀缺性来解释这个矛盾，这种解释也是一种人择原理，因为非常稀有的涨落是观测者出现的必要条件：

> 然而，这种情况的逻辑是，观测者总是发现自己生活在一个能够产生生命的宇宙中，而且这样的宇宙一定非常宽广。我们不可能看到一个寿命低于1010年的宇宙，不然智人就无法演化出来了。

如果要进一步完善这个想法，我们就得考虑宇宙诞生时最可能拥有什么样的性质。假设宇宙的大小是有限的。这样的宇宙还有很多，尺寸各不相同，我们已经知道为什么自己必然生活在其中一个大尺寸的宇宙中，尽管宇宙初始的尺寸可能很小，而且经历了暴胀。另一个我们必须考虑的问题是，满足这些条件的宇宙拥有什么样的拓扑性质呢？

我们已经知道，爱因斯坦方程组能够描述平坦的或者弯曲的宇宙空间。宇宙中的物质分布决定了空间的局部形状，但这并不能告诉你空间整体形状的全部信息。有一个全局的性质还没有被爱因斯坦方程组确定下来，我们必须假设这种性质拥有某种简单的（或者更复杂）的形式。这种性质叫做宇宙的“拓扑”。1922 年，亚历山大·弗里德曼在发现第一个曲率为正和曲率为负的膨胀宇宙解时，就立刻意识到了拓扑性质的重要性。

拓扑性质和几何性质有所不同。在第 2章中，我们讲过弯曲的表面遵循的是非欧几何。让我们把这个想法稍微扩展一下。为了确定一个曲面的曲率，我们在上面标记三个不同的点，A点、B点和 C点。将三个点分别用曲面上存在的最短路径连起来，A到B，B到C，C到A。在一个平面上，我们就得到了一个三角形，内角和是180度。在一个球面上，三角形的边由弯曲的弧线组成，内角和大于 180度，于是这个表面就是正曲率的。在马鞍的表面，三角形的边向内弯曲，内角和就小于180度，于是这个表面就是负曲率的。现在，我们考虑一个平面，其中三角形的内角和等于 180度。将三角形画在一张纸上。把纸卷起来，形成一个圆柱体，让三角形露在外面。仔细观察圆柱体表面的三角形，你会发现它的三条边仍然是直线，内角和仍然是180度。

奇怪吧，圆柱体表面的几何竟然不是弯曲的。能区别柱面和平面的性质就是它们的总体拓扑性质。用这种方法改变拓扑，就可能产生戏剧性的效果。对于宇宙来说，如果它拥有平坦的欧几里得几何，沿着各个维度无限延伸，那么它的体积就是无限的。但如果它卷成一个圆柱并头尾相接，形成一个三维的环状甜甜圈，虽然其中的空间处处都是平坦的，但此时的体积却是有限大的（图11.1）。这就是为什么正曲率空间的体积是有限大的原因。但弗里德曼非常谨慎，没有说平坦的和负曲率的空间体积就是无限大的。他以圆柱体为例，指出将空间的每个维度都黏合起来，就能造出一个体积有限但空间可以是平坦的甚至是负曲率的宇宙。弗里德曼深知这种数学上的精妙之处，但当时的许多天文学家对此并不了解。尽管爱因斯坦的理论拥有这样一种美妙的性质，即宇宙中物质和能量的分布决定了各个地方的空间几何和时间的流逝速率，可是这个理论并没有确定空间的拓扑。[7]

如果除了简单的环状甜甜圈之外还有各种形式的黏合拓扑（wrap-around topologies），而我们宇宙的拓扑又属于其中之一的话，那么即使宇宙实际的大小有限，我们在测量曲率和膨胀速率时，似乎也会认为宇宙是“开放的”，体积无限大，并且会永远膨胀。20世纪70年代，一些宇宙学家研究了这种可能性。1971年，乔治·伊利斯当时在剑桥分类研究了宇宙各种可能的拓扑[8]，结果发现其中一些拓扑会产生与物理学相悖的古怪结论（例如，有些拓扑要求空间两头“黏合”之前要先打个结，于是一个右手①的基本粒子穿过之后就会变成左手的粒子），不过，仍有大量拓扑能够避免其中所有此类的问题。[9]



图11.1 二维空间组成的宇宙也可以像环状甜甜圈的表面一样

① 右手和左手指的是粒子的手征性。例如，伸出右手，四指握拳，指向粒子自旋的方向；拇指伸直，指向粒子运动的方向，那么该粒子的手征性就是右手的。反之，就是左手的。在通常的理论中，某些粒子的手征性是不变的，因此，这种允许手征性可以变化的拓扑就非常古怪。——译者注

如果我们的宇宙呈现为其中一种黏合拓扑，那么天文学家可能会观测什么样的现象呢？1974年，两个苏联科学家德米特里·索科洛夫和维克多·谢瓦茨曼研究了这个问题。[10]他们首先考虑了黏合拓扑会对明亮星系的镜像产生什么样的影响（图11.2）。这就像生活在一个充满镜子的大厅里。每个方向上都存在多个镜像，一个比一个暗淡，因为光绕着宇宙转了一圈又一圈之后才到达我们的眼睛。如果身处一个充满镜子的房间中，你就会在各个方向上看到无数个自己的镜像，一个比一个远，一个比一个小，一个比一个暗。

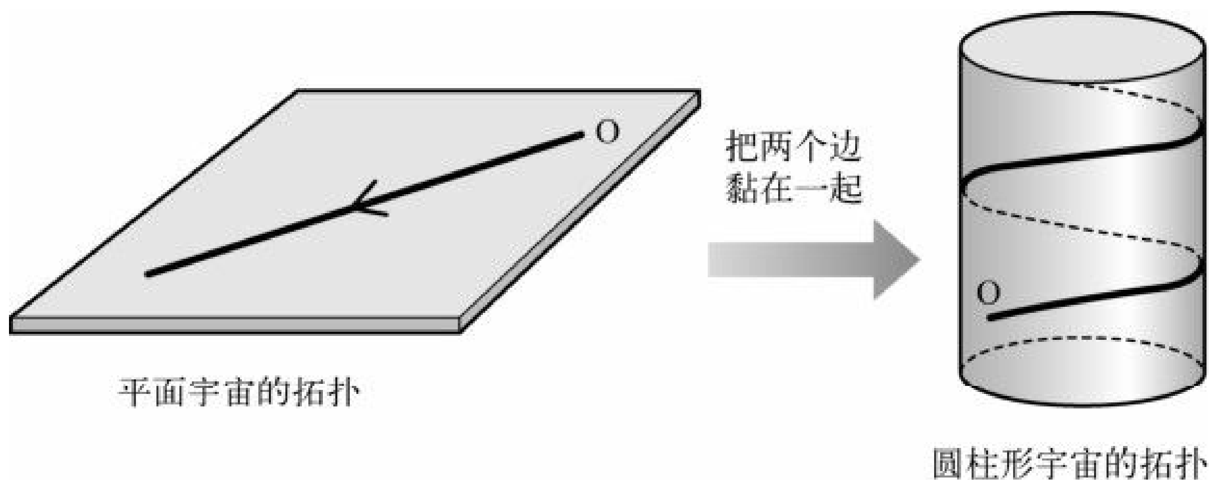


图11.2 将平面黏合成了一个圆柱的拓扑。不像平面宇宙，圆柱形宇宙的观测者可以看到不远处发出的绕宇宙转了好几圈的光子，这些光子都是很久以前发出的

检验自己是否生活在一个黏合拓扑的宇宙中的最简单方法是，选择一个明亮的星系团（巨大的后发座星系团是个不错的选择，它距离我们3.21亿万光年，包含了超过1000个星系），然后在天空中寻找它的镜像。这些镜像互不相同，因为光线沿着不同的路径传过来，镜像也就描绘了星系团不同时期的模样。当然，如果真的存在镜像，第一个镜像应该很明显。

如果宇宙正在利用黏合拓扑制造出体积无穷大的幻觉，那么以上想法就能够对宇宙的实际尺寸的最小值施加一个限制。根据索科洛夫和谢瓦茨曼的分析，以及后来理查德·果特在1980年的复算[11]，整个宇宙的尺寸下限在4～6亿秒差距之间。[12]当时，暴胀宇宙模型还没有出现，主流观点认为，“宇宙”只有一个，要么是有限的，要么是无限的。

1984年，莫斯科的泽尔多维奇和安德烈·斯塔罗宾基回到了从“无”（或者至少从量子真空）中产生宇宙的问题，并假设这时产生的宇宙拥有不同种可能的黏合拓扑。[13]当时，人们认为只有有限的宇宙才能以这种方式产生[14]，所以这些不寻常的拓扑是否比标准的拓扑更容易产生就是个很有意思的问题。他们发现，如果一个简单的平坦宇宙具有有限的尺寸和黏合拓扑，只要三个维度的尺寸都差不多大的话，这种宇宙就有可能像量子涨落一样产生。如果三个维度的尺寸各不相同，宇宙就必须从密度无穷大的奇点中诞生。

仅仅从爱因斯坦方程组出发，并不能确定宇宙的拓扑应该是哪一种。宇宙学家习惯采用最简单的拓扑，但是人们可以争辩说，奇特的黏合拓扑的种类要比简单的拓扑多得多，所以如果我们从魔术师的帽子中

随机抽取宇宙的类型，那么更可能得到的结果应该是黏合拓扑的宇宙。还有人争辩说，有限的宇宙更加自然，不但能避免出现像副本悖论那样的惊人后果，而且能让量子理论变得更简单。这表明，我们必须考虑到所有有限的宇宙，而不仅仅是弗里德曼首先提出的最简单的球对称拓扑。反对黏合拓扑的宇宙的人则怀疑，为什么黏合空间的尺寸必须限定在某个范围以内？这个尺寸在现在刚好可以被观测到，这难道不是个天大的巧合吗？的确，如果宇宙的曲率为零、空间又完全服从欧几里得几何的话，这个反对的理由就非常充分。没有理由能够解释，为什么黏合空间的尺寸应该跟我们的视界的尺寸差不多，其尺寸应该可以任意取值才对。然而，如果宇宙是负曲率的，那么它的曲率半径跟可见宇宙的尺寸非常接近，而黏合空间的尺寸又跟曲率半径密切相关，这样就不算是纯粹的巧合了。

最近几年，天文学家试图解释宇宙微波背景辐射温度涨落中的每一个细节，于是这个拓扑的问题再度流行起来。起初，人们认为微波背景辐射中角度相差十度的温度涨落强度略有不足。对此黏合拓扑的宇宙能作出解释，因为对于一个“小尺寸”的有限宇宙，测量温度波动的大尺寸有时就会容纳不下了，当测量温度波动的尺寸接近黏合拓扑的宇宙的尺寸时，涨落就会趋于消失。[15]然而，最近有一些证据表明，其中一些所谓的不足可能并不具备统计学意义。毫无疑问，在未来的几年里，我们会听到有关这个问题的更多消息，因为从2011年起，我们将得到普朗克卫星的观测数据。

量子的宇宙

“您拨打的号码不存在，请将您的电话旋转90度后再拨。”

——“虚构的”电话应答消息①

① “imaginary number”既可指虚构的电话号码，也可指虚数。在复平面中，一个纯虚数绕着原点旋转 90度就变成了实数（real number），而“real number”又可指真实的电话号码。——译者注

现代物理学面临的一个最严峻的挑战是，如何将爱因斯坦的引力理论与描述物质和光的量子理论结合在一起。传统上，量子理论支配原子和它们的组分所构成的小尺度世界，此时的引力过于微弱，所以被忽略不计。引力理论支配今天宇宙的大尺度结构。但是如果回溯宇宙诞生的时刻，我们就会遭遇引力和量子理论必须结合在一起的时期，此时，量子世界的不确定性已经延伸到了空间和时间本身的结构之中。值得注意的是，两者分离的时刻是由自然常数定义的。

我们从量子理论学到的东西是，一切粒子都有量子波动的一面。粒子的量子效应波长反比于粒子的质量。如果量子波长超过粒子的物理尺

寸，量子的波动就表现为粒子的内在属性。相反，像你和我这样的庞然大物在顺着街道散步时，量子的波动就可以忽略不计了。设想宇宙等价于一个半径等于宇宙年龄乘以光速的圆球，如果我们计算它的质量，就会知道这些物质对应的量子波长会在什么时候超过球的直径。那时，整个宇宙中有因果联系的区域都表现为一团量子的波动，此时我们就认为爱因斯坦的宇宙学方程组不管用了。这个转折的时期非常早，大约是$t_Q=10^{-43}$秒。那时，光信号只传播了10^{-33}厘米。

这个时间的“滴答”小得异乎寻常，却又非常重要。这就是时间的自然单位，它由物理定律决定，不受人类偏好的影响。定义这个单位的自然常数只跟宇宙的量子性质、相对论性质和引力的性质有关。[16]这个单位看起来太小了，是因为我们在用“人类”的单位（秒）来衡量它，而秒的大小表征了日常经验中我们能够分辨的瞬间。我们说宇宙的年龄大约是140亿年，也就是说，经历了10^{60}次这样的量子滴答。在这个意义上讲，宇宙已经非常老了。只经历过些许量子滴答的宇宙才称得上年轻。

当宇宙的年龄只有一个“滴答”时，光信号传输所画出的圆球半径只有10^{-33}厘米。这个距离小得无法想象，不过有一个很好的方法能够让我们对它有更形象的了解。拿出一张A4纸，想象你可以将它裁成两半，然后反复裁下去。当你将它减半30次以后，差不多就成一个原子的大小了。减半47次时，就成了一个质子的大小。减半114次时，就成了10^{-33}厘米的大小。只裁了114次就能从一张A4纸得到最小的尺寸，在这个尺寸上，距离的概念仍然具有物理的意义。沿着另一个方向进行，如果你将纸张的尺寸翻倍90次，就会得到今天整个可观测宇宙的尺寸——大约140 亿光年。所以，这些无法想象的距离，大的也好小的也好，实际上都不过是纸老虎。

我们注意到，宇宙诞生后又经过了1亿个量子“滴答”，大约是10^{-35}秒时，才开始暴胀。因此，量子引力的问题并不影响暴胀之类的事情。然而，如果我们想要探索宇宙暴胀之前的样子，追问它究竟有没有开端，就不得不面对悬在头上的量子引力问题。

我们已经知道，20世纪60年代中期，彭罗斯和霍金开启了人们对爱因斯坦方程组的数学性质的研究，他们列出了宇宙经历开天辟地所需的精确条件。可惜，如果我们向前反推量子引力时期，这个结论所需的初始假定就不再成立了。爱因斯坦的方程组可能需要修改，而引力有可能不再是吸引的了。如果存在暴胀，并不意味着宇宙就没有开端，而是说我们无法确定宇宙必然存在一个开端。我们的定理此时就没什么作用了。实际上，驱动暴胀所需的那种物质（我们会在下一章看到）能够解释宇宙现在的膨胀状态，而且必然违反了证明宇宙存在开端的定理所需的假设。

我们已经说过，永恒暴胀的宇宙将我们带入了一个全新的视角，使得我们能够从全局的高度看待宇宙开端的问题。多重宇宙的每一个角落，比如我们的可见宇宙这样的，都有各自的开端。但自我繁殖的过程孕育了数不尽的暴胀宇宙，从整体上讲，无限多重宇宙并没有开端。如果这种理论正确的话，多重宇宙可能就是一直存在的，而且未来将永远存在下去。

关于量子宇宙学，科学家并没有形成共识。这个理论从许多方面来说存在一定的问题，因为我们习惯于利用量子力学来预言观测者在测量作出时所能看到的结果。比如说，在已知的能量范围内，预言会有多少电子从放射性衰变中产生。在测量作出之前，一切事物都是不同概率的波动；测量作出之后，就记录下一个确定的结果。我们所做的事情，就是预言每一种结果出现的概率。然而，对宇宙来说，并没有开展测量活动的外部观测者，于是量子力学的全部哲学就发生了变化。我们必须预言不同事物之间的关联性。例如，如果今天我们观测到宇宙的膨胀速率等于某个数值，那么我们测出星系团延伸到某个尺寸之外的概率应该是多少？但是在实践中，我们现有的能力还远远做不到这一点。

许多人尝试用一种特殊的方程来研究宇宙的量子力学，这种方程是1967年由约翰·惠勒和布赖斯·德威特（1923～2004）发现的，叫做“惠勒–德威特方程”（德威特总是管它叫爱因斯坦–薛定谔方程，而且把它归功于惠勒，而惠勒则坚持管它叫德威特方程。1988年，他们最终达成一致意见，管它叫惠勒–德威特方程）。[17]

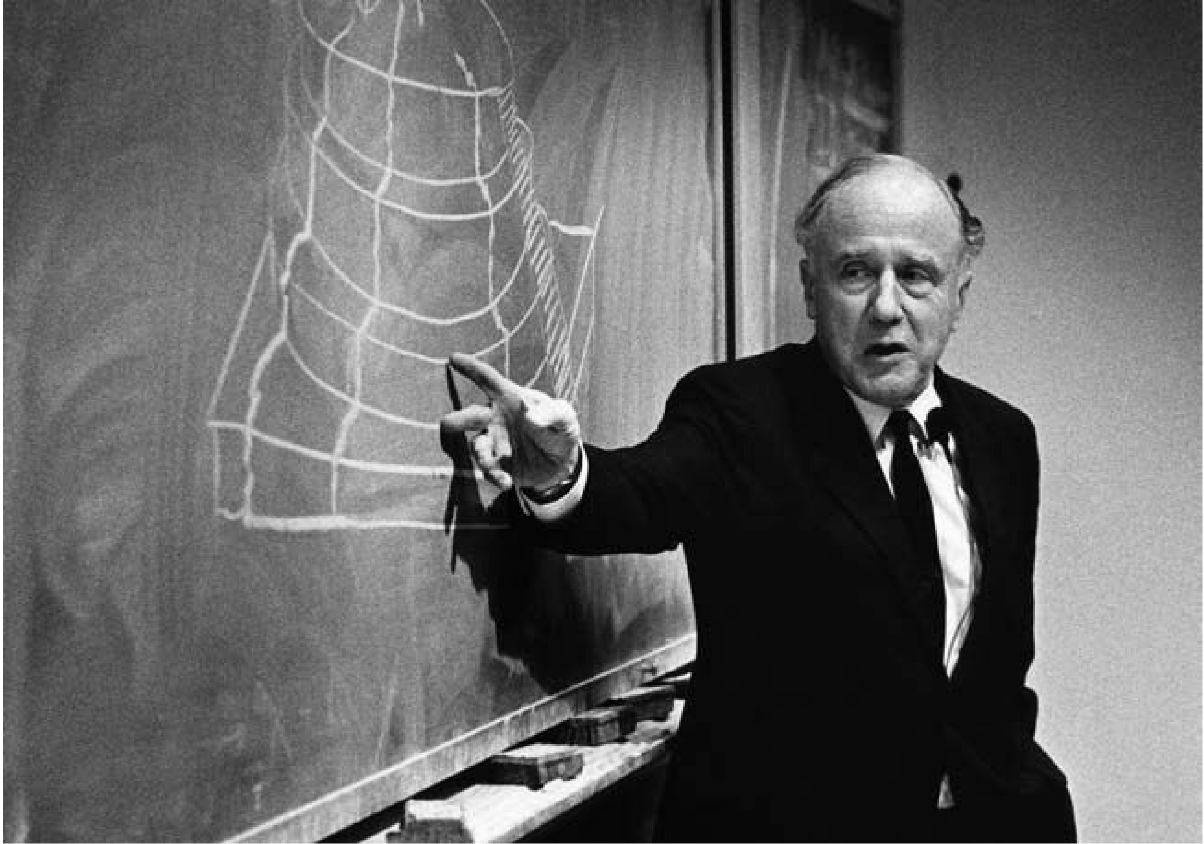


图11.3 约翰·惠勒

惠勒–德威特方程第一次尝试将爱因斯坦的广义相对论方程与描述量子的波函数随着空间和时间变化的薛定谔方程结合在一起。这个方程的解就是宇宙的波函数。如果有人能解出这个方程，就能得知宇宙从一个状态演化到另一个状态的概率。为了找到这样的解，我们又必须确定宇宙波函数的初始条件，但我们还不清楚这些初始条件到底是什么。

对于这个问题，詹姆斯·哈特尔和史蒂芬·霍金的方法非常激进。[18] 1982年，霍金在梵蒂冈教皇科学院的一次演讲中概括了其中的想法。[19] 对于宇宙的“开端”，梵蒂冈有着特殊的兴趣，这似乎与乔治·勒梅特产生了共鸣，他是大爆炸宇宙模型的创始人之一，并在 1960～1966 年间担任过教皇科学院的院长。哈特尔和霍金运用了理查德·费曼发明的一种量子力学的优美形式，计算了观测者发现宇宙处于某种特定状态的概率。



图11.4 1963年，布赖斯·德威特和塞西尔·德威特（Cecile DeWitt）在法国的阿尔卑斯山

为了计算宇宙从状态A演化到状态B的概率，你必须考虑时空中从A到B所有可能的演化路径的影响。在量子力学中，宇宙通过其中任何一条路径都有相应的概率，但是当宇宙变得越来越大，越来越老时，量子效应就越来越弱，越来越不重要，其中的一条路径就占据了主导地位，而其他路径的影响就相互抵消，就像两列波的波峰和波谷相互叠加一样。这条主要的演化路径叫作“经典路径”，对应于爱因斯坦，甚至牛顿，在不考虑量子力学的情况下给出的结果（图11.5）。

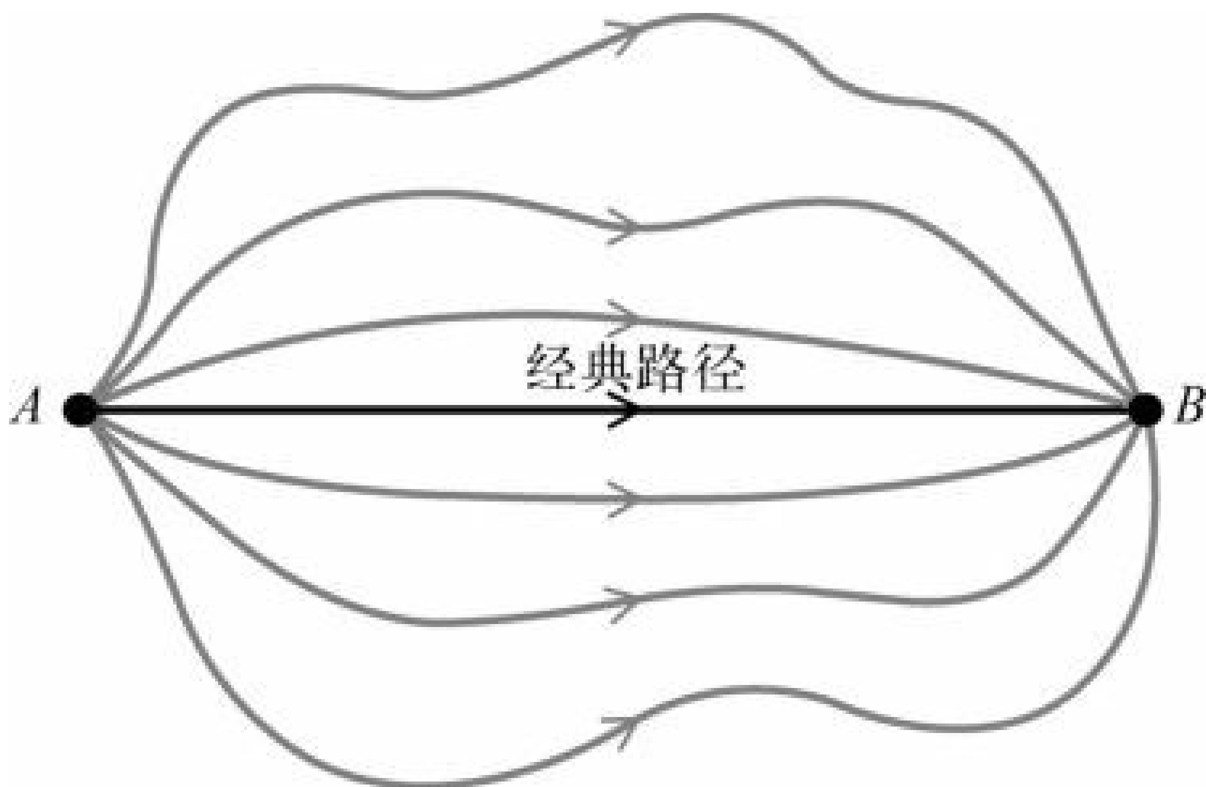


图11.5 宇宙在 A和B两个状态之间的演化路径。牛顿定律认为，宇宙会沿着一个“经典路径”演化。量子力学只能告诉我们宇宙从 A 状态转为B状态的概率，这个概率对A和B之间所有可能的路径都进行了加权平均，图中画出了其中一部分路径

通常，在人们的计算结果中，那些可能的演化路径只包括时空中从A到 B的“洛伦兹”路径。这是一种由光速或亚光速运动的粒子所画出的常规路径。你走路或骑车上班的时候，就在沿着时空中的一条洛伦兹路径运动。但是哈特尔和霍金想要包括另外一种路径，叫做“欧几里得”路径，此时，时间被转化成了空间的第四个维度。这听起来很奇怪，不过物理学家计算这类问题的时候早已将时间转化成了空间，因为这样能够大幅度地简化问题。最后，他们再把其中一维空间变回时间的样子。这是个简便的方法，就像用不同的坐标系画出图形一样。但是哈特尔和霍金并不想把这个变换仅仅看成一种简便的计算技巧：他们提出了一种初始状态，其中的时间也变成另外一维的空间了。

这听起来非常古怪。不过，我们的理论或许可以不要时间。为了标记宇宙不同的状态，区分过去和未来，你不用提到“时间”两字就可以得到所需要的一切信息。甚至在今天，我们也能发现一些类似的情况。如果我们把微波背景辐射的温度当作时钟，那么用下降的温度就可以区分

过去和未来。在哈特尔和霍金的量子宇宙中，当宇宙又小又炽热的时候，从一个状态演化到另外一个状态的路径就由欧几里得路径主导，而当宇宙又大又寒冷时，就由洛伦兹路径主导。

这导致了两个惊人的结论。时间并不是这个理论中的基本概念。时间是一种性质，当宇宙变得足够大，标志性的量子效应变得微不足道时，它就从宇宙中层展（emerge）①了出来：时间是某种只有在非量子的条件下才会具体体现的性质。当哈特尔–霍金的宇宙尺寸很小时，它就会由欧几里得的量子路径主导。时间的概念消失了，宇宙愈发像一个四维的空间。这种宇宙没有时间上的开端，因为时间消失了。

① 有些概念只存在于宏观系统中。例如，一个粒子没有温度，只有当大量粒子达到热平衡状态时才有温度，因此我们可以说温度是一种层展的概念。——译者注

我们通常可以将这种方法理解为时间变成了“虚数”，因为将时间变为空间维度的变换等价于将时间坐标乘以一个虚数因子，即-1的平方根。如果你不是数学家，你会觉得这听起来很离奇，但这是一个十分简单的几何化解释。这对应于将我们图中的时间轴旋转90度，于是变成空间的坐标轴。如果我们观察一个简单的时空图，光线沿着直线从过去射向我们，此时，我们就可以看出虚实的转换对过去产生了什么影响。竖直的时间轴旋转之后变成了水平的，像空间轴一样，而过去就成了一个光滑的圆形表面，就像瓶底一样（图11.6）。

这就是哈特尔和霍金提出的“无边界”状态，是宇宙的起源，而且是1988 年出版的霍金最畅销的书《时间简史》的核心内容。它并没有把宇宙的开端描述成一个划分时间边界的事件。宇宙有一个开端，但并不是大爆炸的奇点，温度和密度并不是无穷大，时间和空间也没有被破坏。它的开端是光滑的、不起眼的，就像路过地球表面的北极极点②一样。[20]

② 这个比喻的大致意思是，北极点上没有北，就像处于“无边界”状态的宇宙没有时间。——译者注

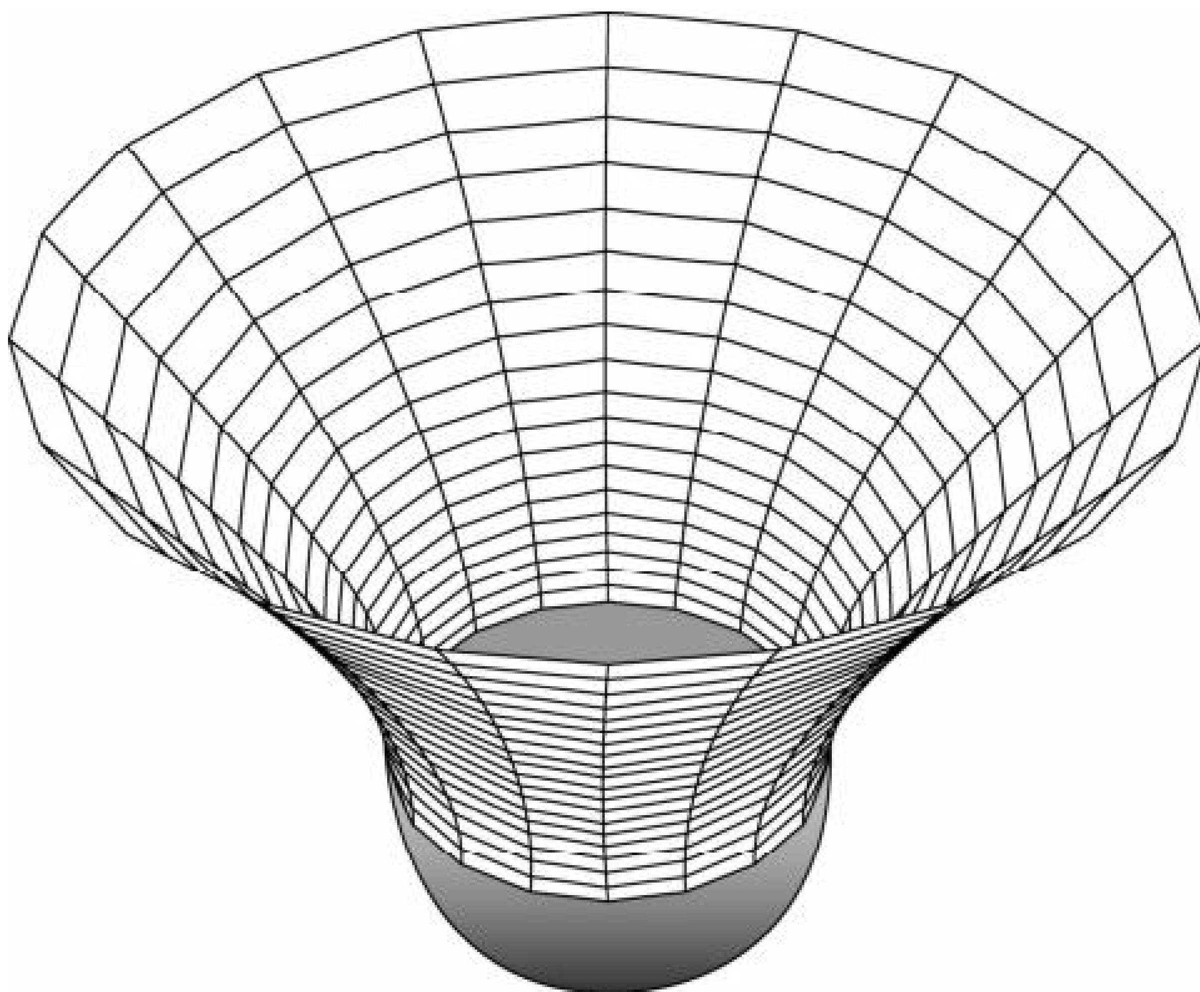


图11.6 在哈特尔和霍金的“无边界”理论中，宇宙初始状态的时间变成了第四维空间（或者叫做“虚时”），当时宇宙的体积非常小。时间轴本应该是上下方向的，但在图中弯曲成了水平方向的，这样就形成了一个光滑的圆形边界，使得时间没有起点。早期宇宙的时空演化就像一个羽毛球

实际上，这个无边界条件想要描述的是，宇宙在一次量子事件中从无中产生[21]时的状态。这种宇宙的故事是这样开头的：很久很久以前，那时还没有时间。

不过，这种表述并非没有问题，它也不是宇宙唯一可能的初始量子状态。艾力克斯·维连金提出了另一种初始状态，对于最有可能从无中产生的宇宙，这种想法得出了完全不同的结果。[22]最终，维连金的初始条件看起来更合理，因为它倾向于产生体积小、温度高、密度大的宇宙——非常像标准宇宙学理论中炽热的早期宇宙。而哈特尔–霍金的表述则要求，最可能产生的宇宙是无穷大的，而且空无一物。[23]

自己制造自己的宇宙

世界上唯一的新事物就是你所不知道的历史。

——哈里·S.杜鲁门（1884～1972，美国前总统）

无需借助高深的虚时和量子力学，我们也能推导出同样的结论，即宇宙的开端并不是炽热的、奇异的，因为当我们向越来越久远的过去追溯时，时间的概念就消失了。我在1986年提出了一个简单的例子，如果要求时空中的所有路径都形成巨大的闭合回路的话，就会产生一种有过去但没有开端的宇宙。[24]

我们在哥德尔的不膨胀宇宙模型中见到过，闭合回路的存在会导致时间旅行，这是爱因斯坦方程组所容许的。那个宇宙不膨胀，但我们也可以设想一个膨胀的宇宙，其中所有粒子和光线的路径都形成了闭合回路，回路的周期极为漫长（超过一千亿年）。在这个宇宙中，我们不会注意到任何诡异之处。但如果我们回溯宇宙的过去，最终却能预知宇宙的未来。这种宇宙并没有时间上的开端。它只是“存在”（is），不过就像比尔·克林顿一样，你必须解释清楚“存在”的含义①。

① 在莱温斯基丑闻中，克林顿曾在否认存在不正当关系时用“there is not”玩弄文字游戏。“there is not”既有不存在不正当关系之意，也可作现在不存在不正当关系解。——译者注

这种模型回避了时间的起源。后来，普林斯顿大学的理查德·果特和李立新在此基础上对细节上加以完善。[25]他们修改了永恒暴胀宇宙模型，提出宇宙也可以自己制造自己。

永恒暴胀的宇宙见证了新的“婴儿”宇宙不断从母宇宙中产生的过程。如果我们生活在其中一个宇宙中，就能追根溯源找出我们的“母”宇宙，以及母宇宙的母宇宙，依此类推。我们已经知道，可能（有人也许会说，是非常可能）追根溯源的过程永远不会有尽头，而制造宇宙的过程或者说多重宇宙并没有开端。但是果特和李提出了一种新的假说，认为其中一个子宇宙的分支会绕回去，自身构成一个回路，从而形成一段闭合的时间循环，看起来就像自己“制造”了自己。如果所有的分支都来自同一个或者几个这样近亲繁殖的分支，宇宙就成了它自己的母亲，因此就不存在所谓的开端了（图11.7）。

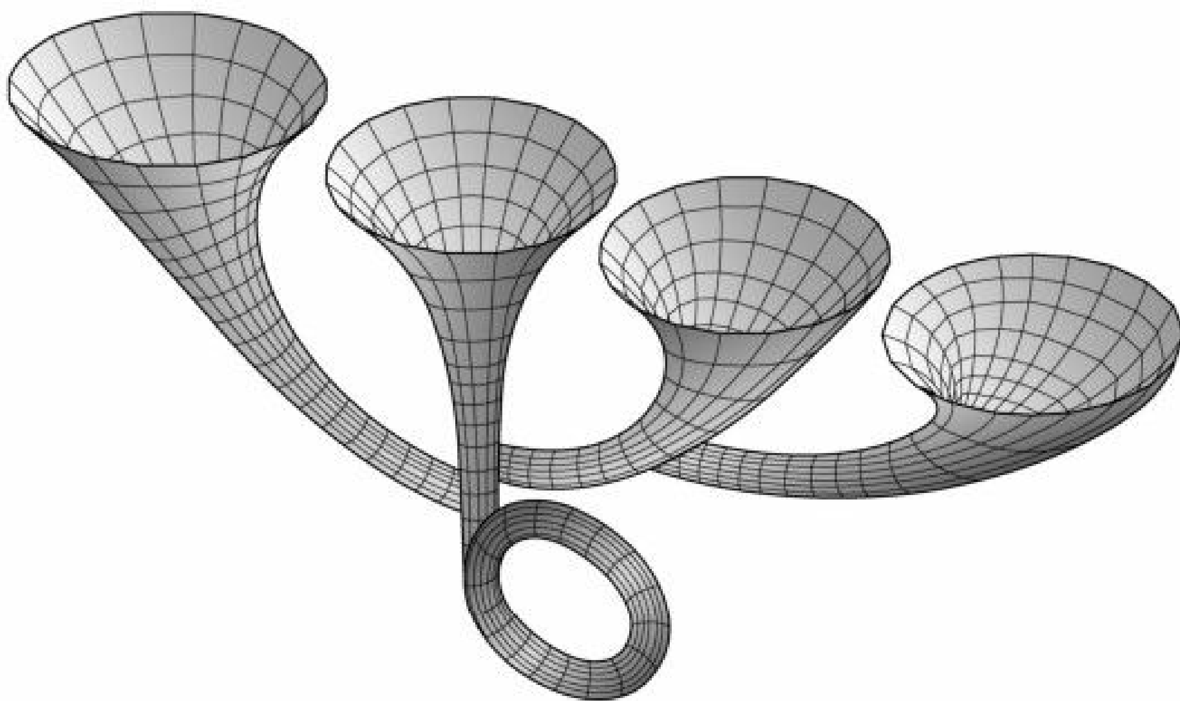


图11.7 果特和李的“自己制造自己”的宇宙。开始时的时间形成闭合回路，因此宇宙就是它自己的祖先。这是哈特尔–霍金模型在非量子力学、“实”时条件下的对应版本

相撞的宇宙

最近我在 M5高速公路上行驶时，我的车被一辆呼啸北去的卡车超过了，车上面写着“马尔文矿泉水”的字样。几分钟以后，我又看到一辆类似的车在向南开，上面写着“高地矿泉水”。现在人们都疯了吗？

——史蒂芬·派米诺夫[26]

从前有个说法是，如果膨胀宇宙的密度超过了临界值，它就会转为收缩，经历大塌缩，然后再“反弹”，回到一个新的膨胀周期，接着再收缩再反弹，永无止境。这个模型是“一次性”闭合宇宙的扩展，是由理查德·托尔曼在1934年首次提出的。但是他又指出，热力学第二定律会导致宇宙的能量不断转化成更无序的形式，例如辐射，这会导致循环的尺寸和周期不断变大。最近，在1995年，马留斯·东布罗夫斯基和我证明了，如果存在爱因斯坦的宇宙学常数，那么不管其数值有多么小，这种振荡总是会结束，最后经历的是指数膨胀（图3.20）。[27]

2001年，循环宇宙模型又重新唤起了人们的兴趣。贾斯汀·考利、伯特·欧伏特、保罗·斯坦哈特和奈尔·图罗克在弦论的基础上提出了一个新的版本。[28]他们称之为“浴火重生”的宇宙（ekpyrotic universe），这是为了纪念古希腊支持循环宇宙观点的斯多葛派哲学家。希腊单词ekpyrosis表示那场毁天灭地的大火，每隔一段时间宇宙就会陷入这场大

火，毁灭，然后像凤凰一样重生。

新版的循环宇宙模型认为，宇宙存在一个拥有最大对称性的天然初始状态。弦理论只有在空间维度超过我们所熟悉的三维时，才有可能成为自洽的包罗万象的理论。为了让这一预言与我们日常的经验相符，人们认为其中只有三个维度变得非常大，也许宇宙的暴胀以某种特殊的方式选择性地作用在了这些维度上，而其他维度至今仍然十分微小，人们对此无法察觉。维度的大小分化必然发生在宇宙非常早的时期，大概在10^{-43}秒左右。浴火重生理论的创始人认为，宇宙有一个天然的初始状态，即两个三维宇宙（叫做“膜世界”[29]）沿着额外的维度相互靠近。这两个宇宙的运动就像两个完全平行的能量箔相互靠近，最后撞在一起。它们相撞时会产生熊熊大火，继而发生反弹，回到膨胀的状态。人们认为，这样的碰撞和反弹并不会产生传统大爆炸模型面临的棘手问题，因为其中并没有产生温度、密度等物理量的无穷大，而且空间和时间的结构仍然是光滑的。碰撞释放的能量转化成了基本粒子，进而影响了宇宙的膨胀。人们希望这个理论所预言的膨胀速率恰好是临界值，而且宇宙的温度分布图之中会表现出微小的涨落，与人造卫星正在进行的探测相吻合。

这个大胆的理论能不能做出一些与暴胀理论不同的可观测的预言呢？这个理论自洽吗？这都是令创始人头疼的关键问题。这个理论可以只碰撞一次，将收缩转化为膨胀，也可以经历无数次的膨胀和收缩。[30]随着一个又一个周期的轮回，宇宙的总熵持续增加，但是你所能观测到的那部分宇宙中的熵并没有大得令人无法接受，因为在每一个周期里，早期急促的加速膨胀稀释了上一个周期中产生的熵。但是这个理论中也有一个与之前类似的问题，那就是随着一个又一个周期的轮回，长寿命的黑洞会逐渐积累，赶上下一次碰撞，并且每次发生收缩以后，都会造成各向异性在新一轮膨胀中急速增长。[31]

其他纯理论的高能物理机制也吸收了膜世界的部分设想，不过并不包括那毁天灭地的碰撞。在空间的另外一个维度中，离我们的宇宙非常近的地方也可能存在另一个三维的宇宙。如果其他的膜世界与我们发生相对运动，它就会改变现有的物理状态，产生可观测的效应：可能会改变传统的自然常数，也可能会激发宇宙的暴胀。不过，研究这种理论的主要动机并不在宇宙学。有人提出，引力相互作用在所有的空间维度中传播，而其余的强、弱和电磁相互作用则不是这样（图11.8）。这可能可以解释为什么引力比其他相互作用弱得多，甚至这还能解释为什么宇宙中的发光物体似乎比参与引力相互作用的物体少得多。从引力传播的距离来看，附近的膜世界可以离我们很近，但从光线传播的距离来看，又会是十分遥远，图11.9是这种情况的示意图。

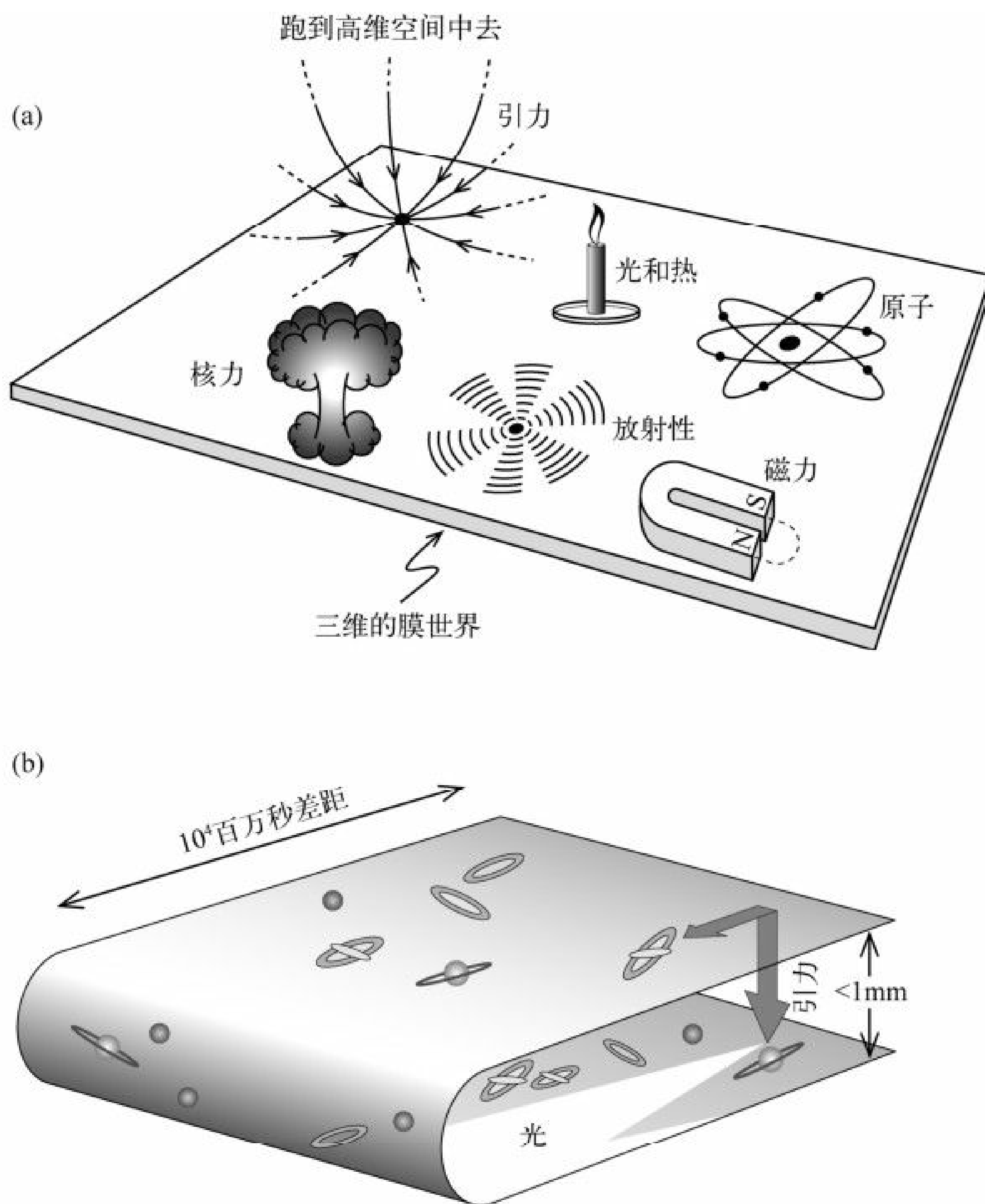


图11.8 (a)电磁的、放射性的和原子核的相互作用被局限在了三维的膜世界之中，而引力则能够在空间的所有维度中传播。因此，引力比膜世界上的其他相互作用弱得多。(b)两个膜可以相互平行，仅仅相距不到 1 毫米，也可以紧贴着自己折叠起来。光线在折叠的膜空间中传播，而引力可以在高维空间中传播，并且能够穿过折叠的膜之间的“空隙”，这个距离要比沿着膜世界到达另一端的距离小得多，因此实际产生的引力效应要比“看”上去应该产生的强得多

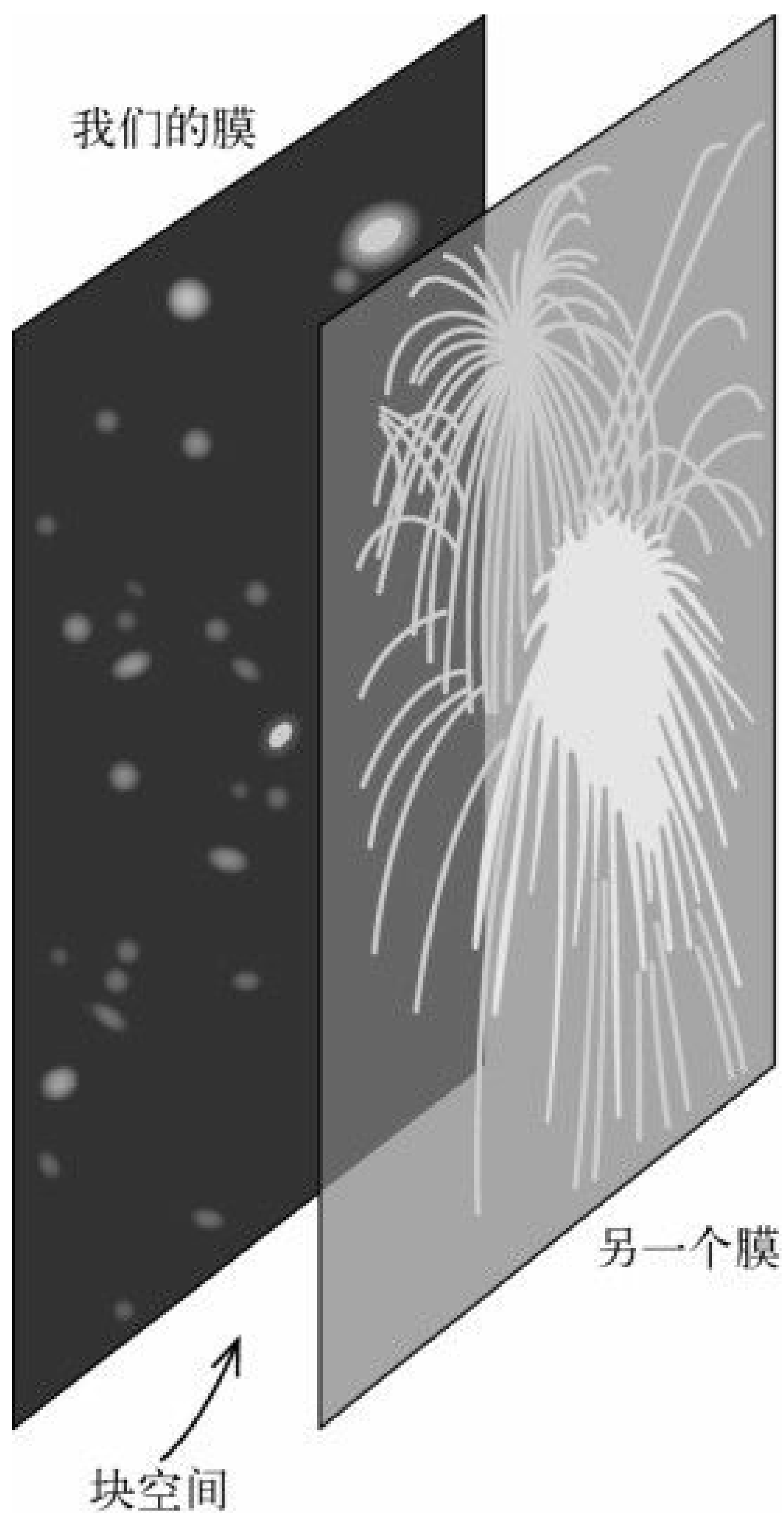


图11.9 我们的宇宙可能处在一个膜上，和不远处另外一个膜保持平行。两个空间位于“块”空间的一个截面上

说起混沌的永恒暴胀宇宙模型时，默认的假设是，我们所在的这一片空间经历了巨大的暴胀，比可见宇宙的范围还要广阔得多。尽管这一点很不自然，但可能是正确的，因为如果暴胀的程度仅仅局限于能够解释我们所看到的东西，那么我们早就受到“隔壁宇宙”不同类型暴胀的影响了。如果真是这样，我们就会看到隔壁存在一个空间泡，并且对宇宙膨胀产生了影响。最终产生的主要影响是，我们会看到宇宙的膨胀变得扭曲了，微波背景辐射的各向同性也被破坏了。不过，如果外来的泡泡第一次和我们的空间发生接触，又会产生什么影响呢？人们对这个问题做了一些尝试，他们研究了一个空间泡经过蜻蜓点水式的接触之后，

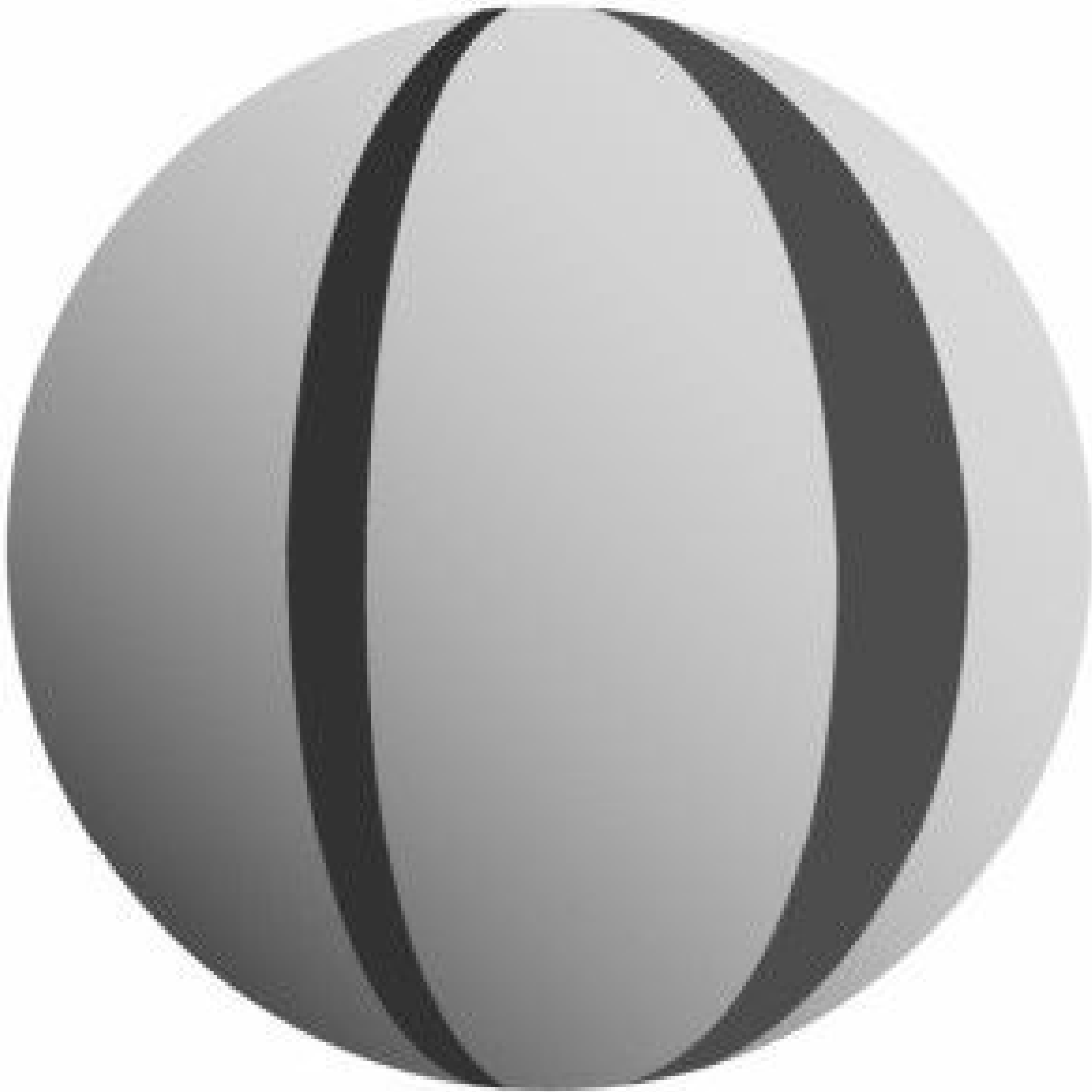


图11.10 两个泡泡“宇宙”相撞以后，会在微波背景辐射中产生温度条纹

在我们宇宙的边缘留下的某种“疤痕”看起来是什么样子的。我们很可能看到天空中的微波背景辐射的温度谱中出现了一些条纹。[32]迈克尔·萨利姆模拟了这种效应，图11.10画出了微波背景辐射中的条纹。不过，这只是其中最简单的一种模型；如果情况更复杂，来犯的时空中可能存在相当不一样的物理机制（或者里面只有反物质），那么它同我们的时空泡接触时就会导致灾难性的后果。幸运的是，我们估计宇宙相撞的概率十分低，不过我们也一度认为小行星和彗星撞击地球的概率也很低。

光之死亡

不要温和地走进那良暮……愤怒吧，为光之死亡而愤怒。

——迪伦·托马斯（1914～1953，英国诗人）[33]

我们知道，在几种宇宙学模型中，传统的自然常数会随着时间变化，甚至还会随着地点变化。安迪·阿尔布雷希特、荣·马戈依卓[34]和我在1998年时提出了其中一种最为激进的模型。这个模型研究的是，如果早期宇宙的光速是可以变化的，又会产生什么影响。这种想法就像短暂的暴胀或者宇宙的加速膨胀一样，会导致一些非常有趣的现象。如果宇宙的甚早期存在一段光速不断降低的短暂时期，宇宙膨胀的速度就会趋近开放宇宙和闭合宇宙之间的临界值，而且磁单极子也会消失，不规则性会被抹平。[35]所有这些现象都令人非常满意，也为当初20世纪80年代宇宙学所面临的诸多问题提供了另一种解决方案。这个模型似乎还能降低宇宙学常数项的影响，而这是暴胀模型无法做到的。[36]

这种非典型的模型不要求今天的光速仍在发生变化。为了产生这些现象，光速的变化只需要像暴胀一样，在宇宙仅仅诞生10^{-35}秒后的一瞬间完成。该模型所面临的挑战在于，它得想办法产生宇宙的密度涨落，这是星系形成的种子，而且所产生的涨落还要体现在微波背景辐射的各向异性之中。暴胀模型可以同时满足这两个要求，因此看起来很诱人。像这样的宇宙学模型，有人尝试用别的方法来产生涨落，不过这些都还需要进一步的研究。[37]

这种光速可以变化的理论被简称为VSL（varying speed of light）宇宙学。这说明宇宙学家变得多么开明，曾经被认为是神圣不可侵犯的基本常数，在他们看来也许是变化的。导致想法解放的原因之一是弦论，因为人们已经知道许多常数可以从弦论中算出来。从这些理论中得到的真空态体现出巨大的多样性，这意味着其中存在数不尽的自然常数的不同“套餐”，它们都会各自产生自洽的宇宙。这些常数看起来就像通常的事物一样，可以遍及所有类型的取值。如果宇宙从一个真空态演化到另一个真空态，这些常数的取值就会变化。自然常数在万事万物中的地位

显著地降低了。另一种关于宇宙的新想法加速了人们态度的转变，进一步降低了“常数”的地位。这种想法要求我们跳出日常经验中的时空维度，在这个范围之外寻找真正的自然常数。

超宇宙

“先生，你是说，”彼得问道，“在这栋房屋里，譬如说，就在附近，到处都有可能有别的世界吗？”

“这是非常可能的，”教授说，他一边摘下眼镜擦擦干净，一边又自言自语，“我真不懂，这些孩子在学校里，到底学了些什么东西？”

——C. S.刘易斯，《纳尼亚传奇：狮子、女巫和魔衣柜》[①]

① 摘自陈良廷的译本。——译者注

为了寻找一个新的万有理论，弦理论家做出了一个了不起的早期发现：只有当空间的维度超过我们日常生活所熟悉的三维时，这样的统一理论才是合理的。存在额外维度的想法会开启非常多的可能性。实际上，弦理论和更深层的“M-理论”还允许额外的时间维度。只有当空间和时间的总维度取到某些特殊值时，通常是十维或十一维[②]，这些理论才能存在。这往往意味着空间是九维或十维的，时间是一维的。然而，这些理论并不要求特定的空间和时间的维度比：如果时间是三维的，空间是七维或八维的，这些理论仍然能够存在。我们假设理论中的时间只有一维，是因为时间的维度超过一维，就会出现许多怪事：不稳定的粒子迅速衰变，能量消失，未来并非完全由现在决定，而现在也不能唯一地确定未来。这很奇怪，但在逻辑上并非不可能，在物理学中也并非不自洽。在一个存在二维时间的宇宙中，很可能无法演化出复杂的生命，不过，如果额外的时间维度非常短暂，就像额外的空间维度非常微小一样，也许就刚好没什么害处了。

② 五种不同的有对偶关系的弦理论要求时空维度是十维，M-理论要求十一维。

——译者注

也存在这样的可能性，这些逻辑上可能存在的宇宙拥有不同的空间和时间的维度比，并各自占据了一种弦景观里的可能真空态。暴胀产生的子宇宙可能拥有不同的时空维度比，在暴胀的过程中，也可能有不同数量的维度变大了。如果真的如此，最终我们发现自己生活在其中一个有三维空间和一维时间的子宇宙中。无论是哪种情况，时空维度比或是变大的维度的数量的选择，都可能完全是随机的，或者是由某种目前未知的机制决定的。

这些基础物理学的研究进展使得我们认真考虑额外维存在的可能性。其中最发人深省的一点是，我们周围叫做“宇宙”的三维空间，只不过是维度更高的世界的一个影子。打个比方，所有被我们称为“自然常数”的物理量，根本不是基本的物理量。真正的常数存在于九维或十维空间中，而我们只不过看到了其中部分常数在我们三维空间中的投影。因此，我们所说的“常数”甚至不必是固定不变的。如果额外的维度在摇动，或者尺寸在缓慢地变化，我们就会发现三维空间中的常数在以相同比率变化。这就是为什么人们对探测自然常数的变动产生了无比浓厚的兴趣，并为此开展了许多天文学观测和高精度的实验的原因。[38]甚至有越来越多的类星体观测数据表明，决定电磁相互作用强度的常数可能发生了微小的变化，在一百亿年间大约变化了百万分之几的大小。[39]传统自然常数的变动为探测额外维的存在提供了绝佳的机会。另外一些物理学家则希望通过分析大型强子对撞机（Large Hadron Collider）的运行结果，能够找到证据支持或否定另一个紧挨着我们的膜宇宙的存在。比如，进入膜宇宙的粒子会衰变，于是人们就会发现有一些能量从空间中神秘消失了。这些纯理论研究进展的最有趣之处在于，它们并不是天马行空无所限制，而是会根据粒子加速器、地面和太空望远镜的实验和观测数据作出相应的修正。

注释

[1] O. Stapledon,*Last and First Men*,Penguin Books, Harmondsworth (1972; first publ. 1930),p.379.

[2] *Business Life*, British Airways magazine, October 2007, p. 62.

[3] 有趣的是，“标量场”在宇宙早期时可能占据了主导地位，由于标量场只有大小没有方向，不能转动，因此它们会导致宇宙处于一种不转动的没有角动量的状态。

[4] E. Tryon, ‘Is the Universe a Vacuum Fluctuation?’,*Nature* 396, 246 (1973).

[5] G. Gamow,*My World Line*, Viking, New York (1970), p. 150.

[6] 一个寿命为Δt、能量为ΔE的涨落一定满足$\Delta t \times \Delta E \approx h$，其中h是自然界的普朗克常数。

[7] 其中存在一些一般性的限制条件，但并没有将拓扑唯一地确定下来。例如，拓扑结构是黏合的，而且体积是有限的，那么比安基和陶伯所有的负曲率模型都必须是各向同性的，不允许产生各向异性的膨胀。参见：J. D. Barrow and H. Kodama,*Class. Quantum Gravity*, 18, 1753 (2001), and *Int. J. Mod. Phys.*D 10, 785 (2001).

[8] G. F. R. Ellis,*Gen. Rel. and Gravitation* 2, 7 (1971).

[9] 如何列出所有可能的负曲率宇宙的种类，这还是个尚待研究的

数学问题，不过平坦宇宙的分类已经搞清楚了。

[10] D. D. Sokolov and V. F Shvartsman, *Sov. Phys. JETP* 39, 196 (1974).

[11] J. R. Gott,*Mon. Not. R. Astron. Soc.* 193, 153 (1980).

[12] 1秒差距大约是3.26光年，相当于大约 19万亿英里或 31万亿千米。

[13] Ya. B. Zeldovich and A. A. Starobinsky,*Sov. Astron. Lett.* 10, 135 (1984).

[14]在暴胀模型中，人们后来发现暴胀可以产生无限大的开放宇宙。实际上，这样的宇宙有无穷多个。参见：A. Linde,*Phys. Rev.* D 58, 083514 (1998), and S. W. Hawking and N. T. Turok,*Phys. Lett.* B 425, 25 (1998).

[15] J.-P. Luminet,*The Wraparound Universe,* A. K. Peters, Wellesley, Mass. (2008), and J. Levin,*How the Universe Got Its Spots,* Weidenfeld and Nicholson, London (2002).

[16]这个时间的基本单位叫做普朗克时间，由普朗克在1899年定义。更细节的知识参见：J. D. Barrow,*The Constants of Nature,* Jonathan Cape, London (2002).普朗克时间由引力常数G、真空中的光速c和普朗克常数h定义，只有一种方法能将这三个常数组合成量纲为时间的量，即 $t_Q = \sqrt{Gh/c^5}$ 。

[17] B. S. DeWitt,Phys. Rev. 160, 1113 (1967); J. A. Wheeler, ‘Superspace and the Nature of Quantum Geometrodynamics’, in C. D. DeWitt and J. W. Wheeler (eds.),*Batelles Rencontres: 1967 Lectures in Mathematics and Physics,* Benjamin, New York (1968), p. 242.

[18] J. Hartle and S. W. Hawking,*Phys. Rev.* D 28, 2960 (1983).

[19] S. W. Hawking, in H. A. Bruck, G. V. Coyne and M. S. Longair (eds.),*Astrophysical Cosmology,* Pontifical Academy, Vatican (1982).

[20]起初，人们认为这种量子宇宙的空间必须是有限的。但是没过多久，霍金和剑桥的奈尔·图罗克（Neil Turok）就发现这种宇宙的体积也可以是无限大的。

[21]有时候，有人将这种过程描述为始于虚无的量子隧穿效应。某些在牛顿力学中不能发生的物理过程，在量子力学中可能是可以发生的。这叫做量子隧穿。这类似于，如果你的能量不够大，无法翻越一座山的话，可以从中间挖一个隧道，这样也能到达山的另外一端。

[22]情况曾一度不甚明朗，因为维连金在最初的理论中犯了一个错误，结果导致他的边界条件给出了跟哈特尔和霍金相同的结果。但没过多久，一些人注意到了这个情况，维连金在 1984年时将这个错误改了

过来，参见：*Nucl. Phys.* B 252, 141 (1985).关于维连金对这一系列事件的评论，可参见：A. Vilenkin,*Many Worlds in One*, Hill and Wang, New York (2006).这样这两种理论之间存在的真正差异才显露无遗。

[23] A. Vilenkin,*Phys. Lett. B* 117, 25 (1982).

[24] J. D. Barrow, in B. Carter and J. Hartle (eds.),*Gravitation in Astrophysics*, Plenum, NATO Physics series B, vol. 156, p. 240.

[25] J. R. Gott and L.-X. Li, 'Can the Universe Create itself ?',*Phys. Rev.* D 58, 023501 (1998).

[26] Letter to the Editor,*The Independent* newspaper, 8 July 2004, p. 28.

[27] J. D. Barrow and M. Dąbrowski,*Mon. Not. Roy. Astron. Soc.* 275, 850 (1995).

[28] J. Khoury, B. Ovrut, P. Steinhardt and N. Turok,*Phys. Rev.* D 64, 123522 (2001).对此的通俗表述参见：P. J. Steinhardt and N. Turok,*Endless Universe: Beyond the Big Bang*, Doubleday, New York (2007).

[29]膜（brane）和膜世界（braneworld）的术语是将二维的膜（membrane）拓展到了高维的情况。这个术语引发了无穷的笑料，例如"universes on the brane"（由膜构成的宇宙）、p维膜世界简称p-branes（我们在文中讨论的是p=3的情况），又如果从空间中除去一部分区域，我猜应该叫做"brane surgery"。（在英语中，brane与 brain读音相同。因此，"universes on the brain"意指只存在于脑中的宇宙，pea-brain则指笨蛋，"brain surgery"意为脑外科手术。——译者注）

[30]随着循环次数的增加，这种模型中辐射的熵并没有随之增大。因为每个循环都会经历加速膨胀，结果膨胀对熵的稀释速度超过了收缩时的反作用。

[31]起初人们认为，这些各向异性都会随着宇宙的膨胀而迅速减弱，但当人们考虑到一些粒子不会再发生碰撞时，情况就反了过来。参见：J. D. Barrow and K. Yamamoto,Phys. Rev. D 82,063516 (2010).

[32] M. P. Salem, 'Bands in the Sky from Anisotropic Bubble Collisions', http://arxiv.org/PS_cache/arxiv/pdf/1005/1005.5311v1.pdf.

[33] D. Thomas, *Selected Poems*, ed. W. Davies, J. M. Dent & Sons, London (1974), p. 131.这首诗是诗人写给他父亲的，当时他父亲日薄西山，视力也越来越不济。

[34] A. Albrecht and J. Magueijo,*Phys. Rev.* D 59, 043516 (1999); J. D. Barrow,*Phys. Rev.* D 59, 043515 (1998).更详细的内容参见：João Magueijo,*Faster Than Light: The Story of a Scientific Speculation*, Penguin Books, London (2003).

[35]暴胀理论解决平坦性问题的方式是，让物质密度在随宇宙膨胀

的过程中比曲率部分稀释得慢，所以经过一轮暴胀之后，曲率项就可以忽略不计了。可变光速理论刚好相反：它让曲率项和宇宙学常数项比通常物质稀释得更快，所以经过一轮暴胀之后，宇宙就变得平坦了。

[36] J. D. Barrow and J. Magueijo,*Phys. Lett.* B 447, 246 (1999).

[37] J. Magueijo and J. Noller,*Phys. Rev*. D 81, 043509 (2010).

[38]对此更详细的内容参见我先前的书：J. D. Barrow,*The Constants of Nature*, Jonathan Cape, London (2002).

[39] J. D. Barrow and J. K. Webb, ‘Inconstant Constants’,*Scientific American*, 55–63, June 2005.

第12章 失控的宇宙

宇宙也许就像洛杉矶一样。
三分之一是物质，三分之二是能量。

——罗伯特·克尔什纳（Robert Kirshner，美国天文学家）[1]

最畅销的宇宙

神志正常的人认为自己不被所谓“终极问题”所困扰，而对日常柴米油盐更关心。他们为自己的这种能力感到骄傲，并且称之为分清主次，或者脚踏实地……

——希丽亚·格林（英国心理学家）[2]

1996年夏天，作为建校250周年庆祝活动的一部分，一场盛大的宇宙学会议在新泽西州的普林斯顿大学召开了。[3]当时的天气状况糟透了，又湿又热，还有夹杂着雷电的暴风雨。人们汗流浃背，却发现老式学生宿舍里的空调不过是个摆设而已，起不了什么作用。只有走进大会堂时人们才会得到解脱。这次会议的一个创新之处是，其中不仅举办常规的学术报告，还可以请两个或三个主讲人同台亮相，像政治竞选一样，相互竞争，各自向观众“兜售”自己看好的某个宇宙模型。他们说完自己的广告词以后，就进入了一个关键的辩论环节，观众也可以参加。

总的来说，当时的宇宙学家都对暴胀宇宙模型非常满意。他们不太考虑混沌或永恒的暴胀，多重宇宙这个词还没有收入他们的字典，尽管从以前的人择原理的角度来看，这种观念大家都很熟悉了。会议中讨论最多的议题是一些精确观测的问题，如宇宙的膨胀速率、宇宙的年龄、星系能否及时形成、物质和辐射所体现出来的不均匀性是否能脱胎于某种类型的暴胀在宇宙甚早期所产生的不规则性等。

每一个宇宙学主讲人都要尽可能地证明，他的模型中关于宇宙物质组成和宇宙膨胀行为的描述，与所有的观测结果都吻合得天衣无缝。其中脱颖而出的是麦克·特纳（Michael Turner）的模型，这种宇宙的膨胀速率几乎达到了临界值，这正是暴胀预言的结果，但其中还存在一个取值很小、符号为正的众所周知的宇宙学常数。爱因斯坦首先提出了宇宙学常数，不久之后又抛弃了这种想法。这种宇宙学常数会产生排斥性的引力，导致今天的宇宙加速膨胀。特纳指出，这种模型能够胜出一点也不意外，因为它的大部分内容都与其他竞争对手极为相似，只是多了一个有点儿别扭的物理量（宇宙学常数），这个常数使得它与观测结果符合得更紧密一些。

这个夺得桂冠的模型叫做“Lambda-CDM”模型，其中Lambda（拉姆达）是指宇宙学常数，而CDM是“冷暗物质”（cold dark matter）的缩写。无论从哪个方面来说，宇宙中都应该存在暗物质，因为产生星系和星系团之间强大引力所需的物质比观测到的发光物质的总量大十倍，必须存在许多暗物质才能解释这个矛盾。暗物质非常特殊，只参与引力相

互作用，可能还参与弱相互作用；不然的话，宇宙最初三分钟时所产生的氘核就会减少，与观测结果矛盾。这意味着暗物质最可能的候选者是参与弱相互作用的中微子，或者跟中微子很像的新型粒子。然而，当时已知的各种中微子并不符合要求。1985年，人们第一次用大型计算机模拟了类似中微子的暗物质对宇宙膨胀历史的影响，结果发现它们的质量太小了，所引发的星系结团行为不符合观测的结果。[4]

为了符合所有的观测结果，类似中微子的粒子必须比质子还重得多，这样才会变得很慢，所以是"冷"的——温度只不过是我们给气体分子的平均速度所起的名字。在计算机模拟中，这种慢吞吞的粒子会导致星系的小尺度结团行为有一个显著的特点，与观测结果非常符合。加上了一个新的参数拉姆达后，Lambda-CDM模型的得分就遥遥领先。但是没有人对Lambda-CDM模型的成功而感到欢欣鼓舞，就连模型的提出者也没有。这个模型看起来人工雕琢的痕迹太重了，老实说，很难看。

这个最佳宇宙模型跟勒梅特六十多年前提出的模型很是相像。就像爱因斯坦一样，当时的宇宙学家已经对宇宙学常数失去了兴趣。要想在最佳的宇宙模型中发挥作用，它的数量级必须极其微小（10^{-120}）。这个数实在太小了，以至于许多物理学家都认为，宇宙学常数的真正取值应该是零：物理学中存在一种有待发现的深层机制，能让宇宙学常数的大小正好等于零。总有一天，我们会发现这种新的对称性机制。在此之前就先把宇宙学常数抛在脑后吧。这种态度在粒子物理学家当中很常见。同时，天文学家总是对自己的观测数据不太放心。拉姆达所依赖的基础并不牢固，很有可能会逐渐消失，或者变得比我们起初设想的更不靠谱。甚至连认真考虑Lambda-CDM模型的人都显得格外小心，因为没有找到宇宙学常数存在的直接证据。我们并没有直接看到今天的宇宙在加速膨胀，而是从其他有关宇宙演化的观测结果中找到了一丝迹象。

1998年，戏剧性的转机出现了。由世界顶级天文学家领导的两个大型研究团队分别发现了第一个直接证据：宇宙正在加速膨胀。令人们感到惊讶的是，哈佛大学的亚当·里斯领导的高红移超新星计划和加州大学伯克利分校的劳伦斯伯克利国家实验室的萨尔·波尔马特领导的超新星宇宙学计划都发现了令人惊叹的新证据：从几十亿年之前开始，宇宙就开始加速膨胀了。[5]首先，你需要将哈勃定律扩展到比以往更遥远的距离上，这样才能在很大的范围内比较膨胀速度随距离的变化，看看它究竟像哈勃定律说的那样是速度与距离呈正比，还是速度比距离增长得更快一些。如果曲线向上弯曲，就标志着加速膨胀。

想要精确测定速度的大小其实很简单，只要知道光线的红移量就可以了。但是问题在于如何才能确定你测出的速度所对应的遥远光源的距离。如果一个光源看起来不亮也不暗，这是说明它本身很暗但距离比较

近，还是说明它本身很亮但距离比较远呢？理想的情况是，你恨不得宇宙中的发光物质都是100瓦电灯泡！你可以用望远镜读出每个光源上的100 瓦标签，这样就能知道它们本身的亮度。比较标签对应的亮度和它们的视亮度，你就能推断出每个灯泡距离我们有多远。呵呵，宇宙中并没有这么多贴着标签的灯泡随它一起膨胀。不过，你可以寻找一些绝对亮度恒定不变的天体（就像电灯泡一样）。通过观察它们的某些物理性质，例如亮度的变化率，就可以将它们认出来。这种作为参照标准的天体被天文学家们称为"标准烛光"。

地面望远镜和哈勃太空望远镜能够观测一种特殊类型恒星的爆发，叫做Ia型超新星，而且其中一些超新星距离我们非常遥远。利用这个方法，两个研究团队才得到了新的观测结果。Ia 型超新星是标准烛光的绝佳候选，因为人们认为这种超新星产生于某种特定的宇宙现象，并且是宇宙中最明亮的一群天体。[6]

如果恒星耗尽核燃料后，剩余的质量小于太阳质量的 1.4 倍，它就会在引力的作用下向内坍缩，最后变得跟地球差不多大，这时起支撑作用的是电子和原子挤在一起时所产生的抵抗性压强。[7]这种稳定的状态叫做白矮星，而这样的恒星尸体在宇宙中非常普遍。总有一天，死亡的痛苦会降临在我们的太阳上，最后产生一颗白矮星。

如果恒星残骸的质量大一些，在太阳质量的 1.4倍和 3倍之间，电子的抵抗性压强就无法和引力抗衡了，于是原子就会被压碎。电子会被压进原子核的质子里，变成一堆中子。中子不愿意挤在一起，只要恒星的质量没有超过太阳的 3倍，中子的抵抗性压强就能防止引力的进一步挤压，从而产生一颗稳定的中子星，直径只有区区几千米，密度却比钢铁还要大 100万亿倍。就像白矮星一样，宇宙中也有很多中子星，其中一些自转速度很快的中子星又叫脉冲星，它们自转时会周期性地向我们发出辐射，就像一座灯塔。但如果垂死恒星的质量超过太阳的3倍的话，就没有任何一种已知的自然力能够支撑它了。最终，大量物质将落入一个非常小的区域中，连光都无法逃出来。从外部的宇宙看不到其中仍在进行的坍缩过程：一个黑洞就形成了。

宇宙中大约有一半恒星是成对出现的，并且都在绕着一个公共的质心旋转。如果其中一颗恒星死亡后形成了一颗白矮星，它就会不断从伴星的外层区域汲取物质。最终，这种同类相食的行为会让白矮星的质量超过 1.4 倍太阳质量的极限，于是电子的压强就再也无法抵挡引力的挤压了。这颗白矮星会发生一次剧烈的热核爆炸（图12.1）。只要白矮星的质量一超过 1.4 倍太阳质量，爆炸就必然会发生，而且每颗白矮星爆炸时的峰值亮度都差不多，大约是太阳亮度的10亿倍——这时，一颗恒星变得像整个星系一样耀眼。在爆炸发生后的几个月中，它的亮度和颜

色都会以一种典型的方式变暗。在爆炸发生后的前几天或前几个星期内，这种描述亮度随着时间减弱的“光变曲线”主要由镍元素的放射性衰变决定，然后由钴元素的放射性衰变决定。通过研究峰值亮度和光变曲线的下降速率之间的关系，这两个超新星观测团队算出了不同超新星与我们之间的距离，然后加以比较。

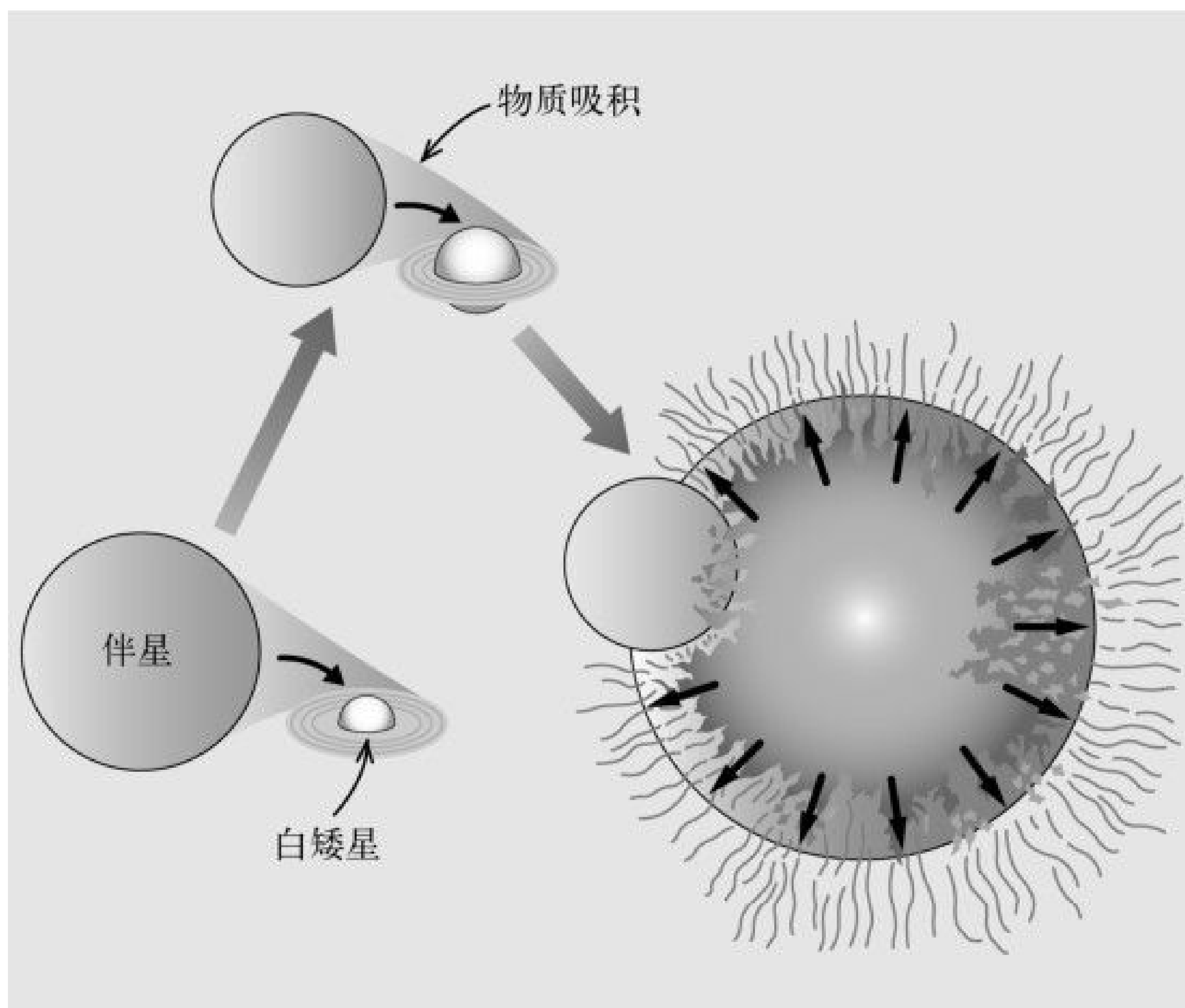


图12.1 当一颗白矮星从它的伴星那里吸积的物质达到一定程度时，就会变成一颗 Ia型超新星。此时，它的质量超过了钱德拉塞卡极限，白矮星无法再支撑自己产生的强大引力。于是白矮星猛烈收缩，随之而来的热核反应引发一场爆炸，我们就看到了一颗超新星

哈佛和伯克利的团队都用这种新的工具来确定超新星的距离，于是便扩大了我们对哈勃定律的测量范围。他们首先用强大的地面望远镜监测夜空的几百个区域，其中每个区域都包含了一千个左右的星系，监测

要在新月的时候进行，因为那时的天空最暗。三周之后，他们又将望远镜对准同样的区域，而且观测得更加仔细，看看恒星的亮度有没有急剧增加，变成超新星。他们一般会找到约二十五颗亮度刚刚开始增加的超新星。然后，他们会用地面和太空望远镜来跟踪超新星的亮度变化，观察它的亮度达到峰值，然后回落到爆发前的水平，同时也在监测超新星颜色的变化（图12.2）。值得注意的是，他们看到的光变曲线形状跟地球附近相同类型超新星的光变曲线形状很相似，这就让观测者们充满了信心。他们相信，在可见宇宙的边缘发现的那些超新星在本质上是一样的，它们的亮度相对暗淡完全是因为它们太远了。

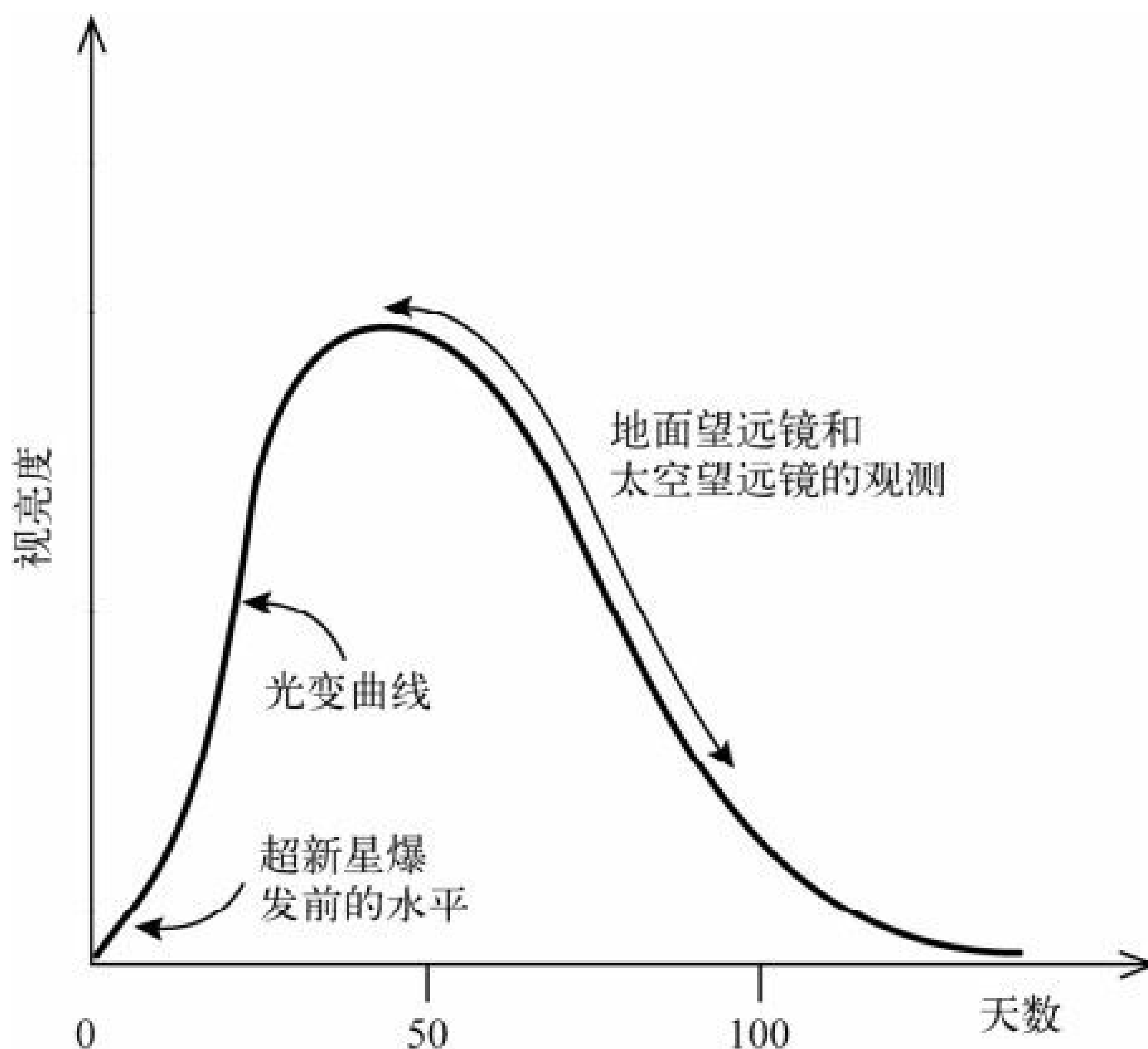


图12.2 Ia型超新星的光变曲线。超新星的视亮度先是达到最大，然后再缓缓回落到爆发之前的水平。在实践中，人们会追踪好几种颜色所对应

的光变曲线

当他们把所有的数据放在一起时，两个小组就得到了同一个结论。哈勃定律的曲线，也就是描述遥远超新星的退行速度同它们的距离之间的变化关系的曲线向上弯了（图12.3）。宇宙在加速膨胀。随着观测数据的日益积累，两个小组用不同的分析技术对不同的数据进行了仔细的研究，又对于其中所有关于标准烛光及其传播空间的假设进行了细致入微的检查，最终这个发现于1998年1月发表，并成为天文学研究有史以来最火爆的热点。

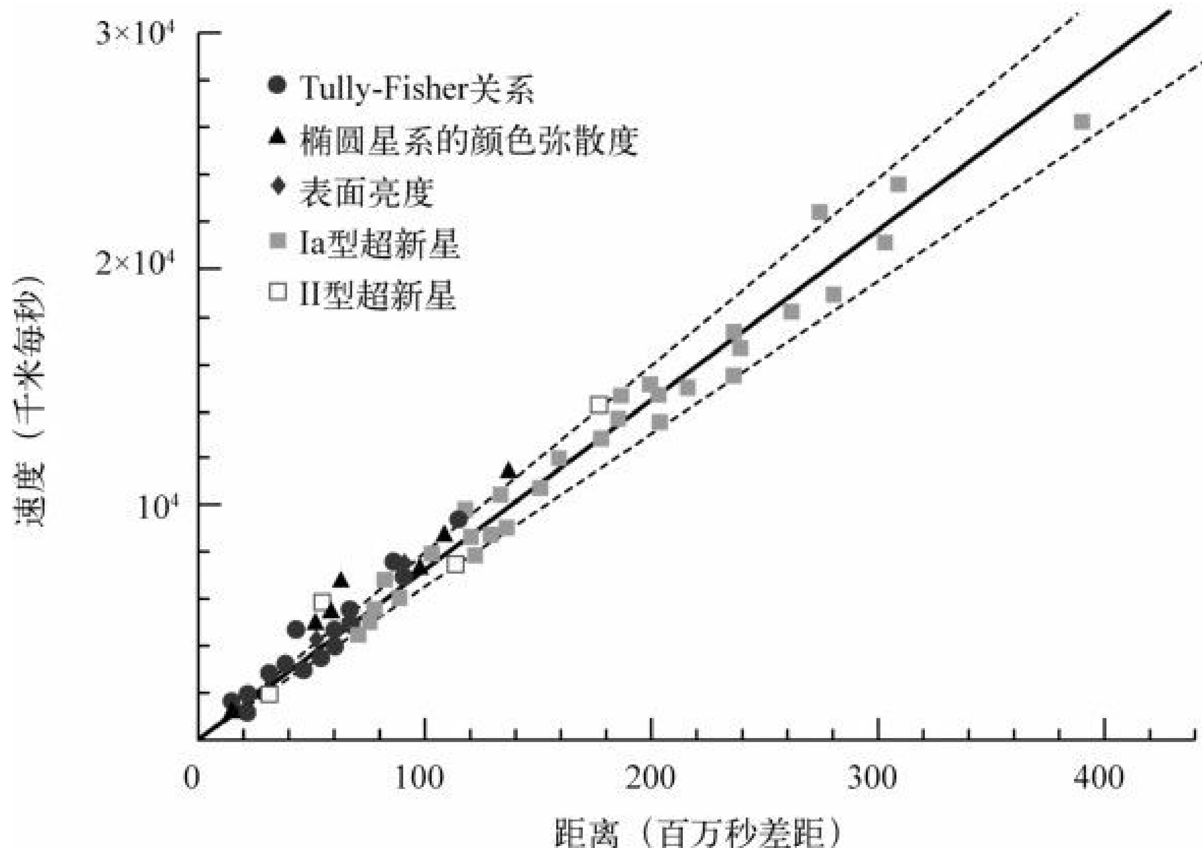


图12.3 退行速度随距离变化的哈勃定律，其中包含了最近的超新星观测数据。这张图说明退行速度随距离变化规律的最佳拟合结果为，哈勃常数等于72千米每秒每百万秒差距

不合常理的宇宙

我们不得不学着理解这个毫无魅力的宇宙，因为我们别无选择……如果没有这么多事实摆在我的面前，你还要跟我讨论这种宇宙的话，要么我会问你是不是嗑过药了，要么我会叫你不要再胡扯了。

——约翰·巴赫恰勒（John Bahcall，1934～2005，美国天文学家）

超新星的观测结果棒极了。这是第一个关于宇宙加速膨胀的直接证据。它非常精确地验证了Lambda-CDM模型的预言，同时也证明了暴胀模型预言的宇宙最初时刻的反引力加速膨胀是可以存在的。关于观测到的加速膨胀，最简单的模型解释得非常成功。如果我们重新拾起爱因斯坦抛弃的宇宙学常数，作为一种产生反引力效应的机制，并假设宇宙的膨胀速率像暴胀所预言的那样，非常接近临界值的话，那么我们就得到了勒梅特曾经提出的一个模型（我们在图 3.13 中见过），而且这个模型与观测结果非常符合。

尽管向爱因斯坦方程组中引入一个小小的宇宙学常数就能出色地解释观测数据，宇宙学家还是觉得事情可能没这么简单。宇宙中可能存在一些反引力的奇异压力，比如那些被认为能够驱动宇宙早期发生暴胀的机制。因此，“暗能量”一词被用来描述这种神秘的能量来源。暗能量的密度可能保持不变（这时它就像宇宙学常数），也可能随着时间变化，最终产生的效应与真正的宇宙学常数在近期历史中的效应类似。值得注意的是，为了符合观测的结果，我们必须要求宇宙中 72%的能量以这种神秘的暗能量形式存在，另外28%的能量才以物质的形式存在。在28%的物质中，只有4%是普通物质，另外24%是某种不是原子物质的冷暗物质。人们还不确定暗物质究竟是什么，不过可能是一种新型的中微子，也许很快就会被欧洲核子研究中心的大型强子对撞机发现（图12.4）。

在过去的十几年中，随着众多间接证据的进一步发现，宇宙加速膨胀的结论得到了越来越多的支持，对驱动加速膨胀所需的暗能量在总能量中所占比例的估计也更为明确。关于遍布宇宙中的微波背景辐射中存在的微小涨落，也不断有各种卫星携带的和地面上的望远镜对此进行了更细致的观测。膨胀的加速会改变回望宇宙时所看到的几何形态（观察过去某个特定时刻的事物时，就会比你设想的更小），于是我们就需要根据观测结果来限制这种变化发生的程度。这也就限制了宇宙所能包含的暗能量和暗物质的总量。最近，第三个限制条件也被人们发现了。宇宙物质密度的涨落最终导致星系和星系团的形成，而在宇宙以辐射为主的时期，这种涨落就应该表现为强烈声波的形式。当宇宙的温度下降到大约三千度时，电子就不再跟辐射发生碰撞了，强烈声波的传播速度也就因此急剧下降。但这些声波还保留着它们刚产生时的尺寸。因此，宇宙中物质的结团模式也应该体现出一些微小的涟漪，在超过 120百万秒差距的尺度上存在一个额外的强烈特征，这就是声波的遗迹。这种效应被冠以一个不雅观的名字“重子声学振荡”（baryon acoustic oscillation，简称 BAO），因为普通物质的原子核是由质子和中子组成的（占总能量的5%），而重子就是这类物质的统称。当天文学观测对准过去某个特定时期的星系时，以上机制就会告诉我们关于星系尺寸的更多信息。

如果将超新星、微波背景辐射和声学振荡的观测结果放在一起看，我们就会发现它们是惊人地一致。这些观测数据的不确定性和确定性恰巧可以相互弥补，所以综合考虑之后，这些观测结果就大大缩小了各种可能性的取值范围。图12.5所示意的结果很明确，其中不确定性的范围（每组数据中由大到小的三个椭圆分别表示正确结果落入其中的概率为99.7%、95%和68%）被不同数据的综合分析大大缩小了。

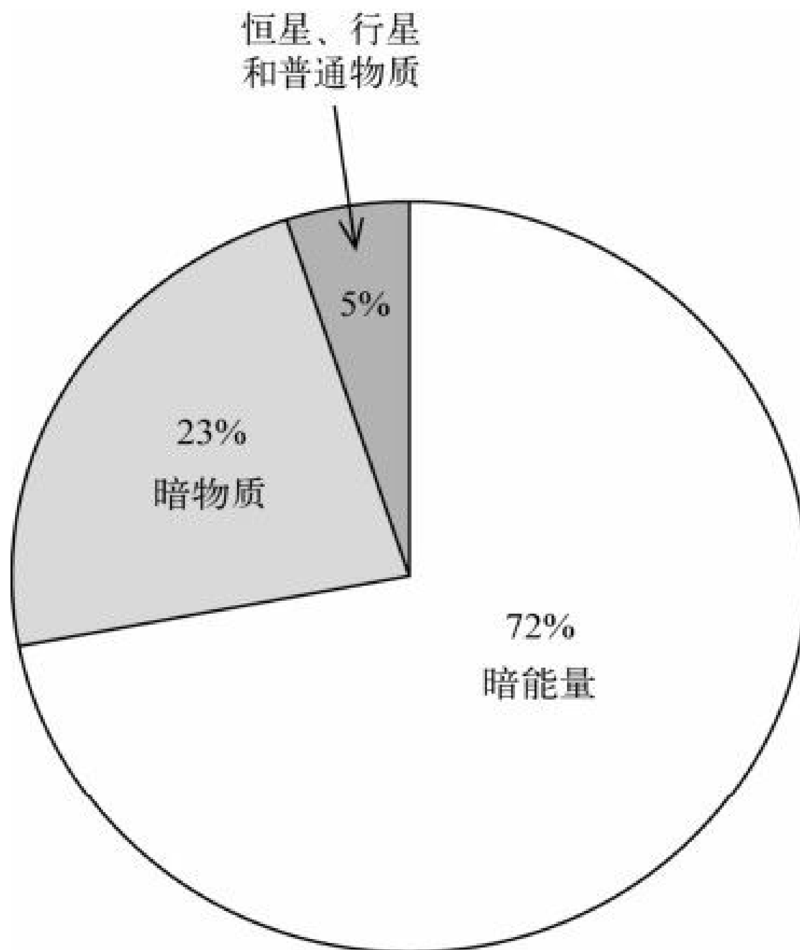


图12.4 宇宙各种物质成分的饼状图。参与引力相互作用的“暗能量”所占的比例最大，它导致宇宙最近开始加速膨胀。其余的部分是暗物质和发光物质。组成暗物质的很可能是某种新型的弱相互作用粒子，很像我们熟知的中微子，不过暗物质粒子的质量要大得多。我们希望在日内瓦运行的大型强子对撞机实验能够证实暗物质的存在，并且能够在地下探测器中直接探测到它们的到来

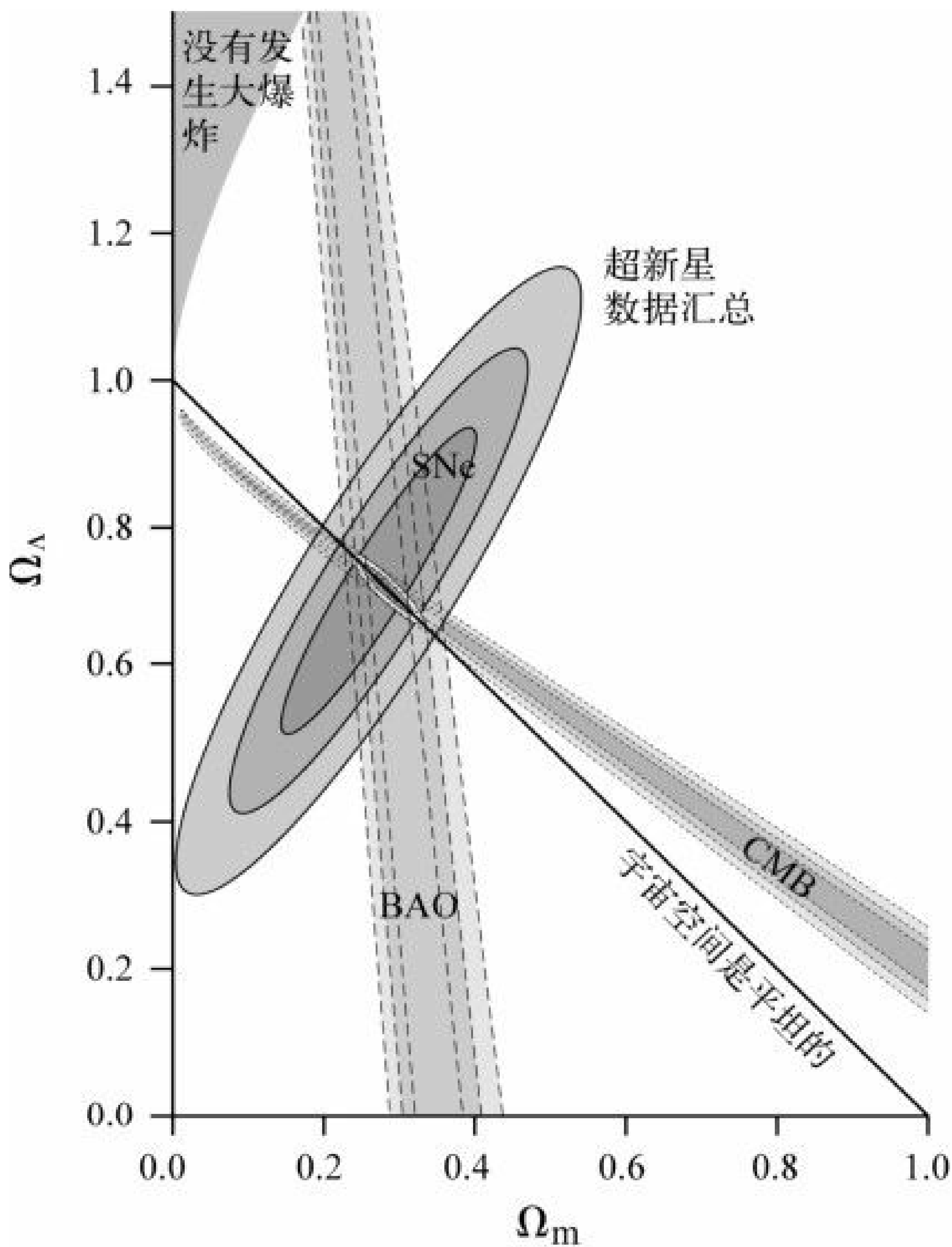


图12.5 图中显示了暗能量占宇宙总能量的比例Ω_Λ和暗物质、发光物质等其他物质所占的比例Ω_m。超新星的所有观测结果（SNe）、微波背景辐

射的涨落（CMB）和物质分布的重子声学振荡效应（BAO）都对宇宙的组成施加了限制。三种明暗不同的区域分别表示 68%（最暗的）、95%（亮度居中的）和99.7%（最亮的）统计置信度。重叠的区域表示暗能量占宇宙总能量的72%，如图12.4所示

垂直的坐标轴表示宇宙中暗能量所占的比例，水平的坐标轴表示其他物质所占的比例。所有数据重叠的狭小区域，对应两种组分的比例分别大致为0.72和0.28。[8]同时，这些数据也集中在曲率接近零的平坦宇宙的范围里，而宇宙空间的平坦性正是我们期望从暴胀中得到的结果。

这个图景是基于一个假设，即暗能量就是爱因斯坦的宇宙学常数。这意味着暗能量的密度始终保持不变，垂直坐标轴则给出了暗能量密度占宇宙总能量密度的比例的可能取值范围。但是暗能量可能比我们预计的更奇特，就像其他形式的物质一样，它的密度会随着时间的流逝而略微有些变化。发生这种情况的条件是，暗能量的压强与其能量密度的比值[9]不等于-1。当这个比值w等于-1时，暗能量就恒定不变，与宇宙学常数的行为一模一样，如图 12.6所示。如果这个比值偏离了-1，暗能量就一定会随着时间缓慢变化。假设我们不排除这样的可能，并再一次利用我们全部的数据，看看所有数据在68%、95%和99.7%的置信度下的重叠区域。垂直坐标表示压强和密度的比值w，水平坐标表示除暗能量以外其他物质占宇宙总能量的比例。如果暗能量就是宇宙学常数的初始假设成立，所有数据都应该汇聚到w等于-1且物质比例为0.28的那一点上。

目前，人们对全部数据的最优估计是，暗能量的压强与能量的比值有 95%的概率位于-1.097和-0.986之间。最简单的情况是 w=-1的宇宙学常数模型，并且现有的数据非常有力地支持了这个模型。于是人们经常假设暗能量就是这种简单的常数。

从许多细节问题来看，我们今天所面临的处境都非常简单，但从总体上看又变得神秘莫测。可见宇宙，可能是无限的多重宇宙的一个无穷小部分，正在加速膨胀。它所遵循的演化路径正好符合乔治·勒梅特在1934年提出的简单模型。[10]空间的几何非常接近平坦的欧几里得几何，而且宇宙似乎正在驶向永远膨胀的结局。想要很好地描述加速膨胀，仅仅加入一个宇宙学常数就行了。爱因斯坦最先引入了宇宙学常数，后来又抛弃了它；然后在1934年，勒梅特又第一个将它重拾了起来。现在我们将宇宙学常数理解为宇宙的真空能。从数据中得到的最佳拟合结果是，宇宙能量密度的 72%是由这种产生排斥性引力的真空形式组成的，而剩下的 28%是由产生引力的暗物质和发光物质组成的。一个简单的数学公式可以描述宇宙的膨胀速率是如何随时间变化的。[11]我们在图3.13

中见到过这种独特的形式。宇宙早期的膨胀行为不受真空能的影响，因为当时真空能的大小可以忽略不计。随着时间的推移，宇宙逐渐开始转型，当它的尺寸只有今天的 57%时，它就开始从减速膨胀变为加速膨胀。不久以后，其他物质的能量密度下降到比真空的能量密度还低，此时宇宙的尺寸大约是现在的73%，大概发生在45亿年前——非常接近地球形成的时间，尽管没有人指出这不应该仅仅是个巧合。[12]

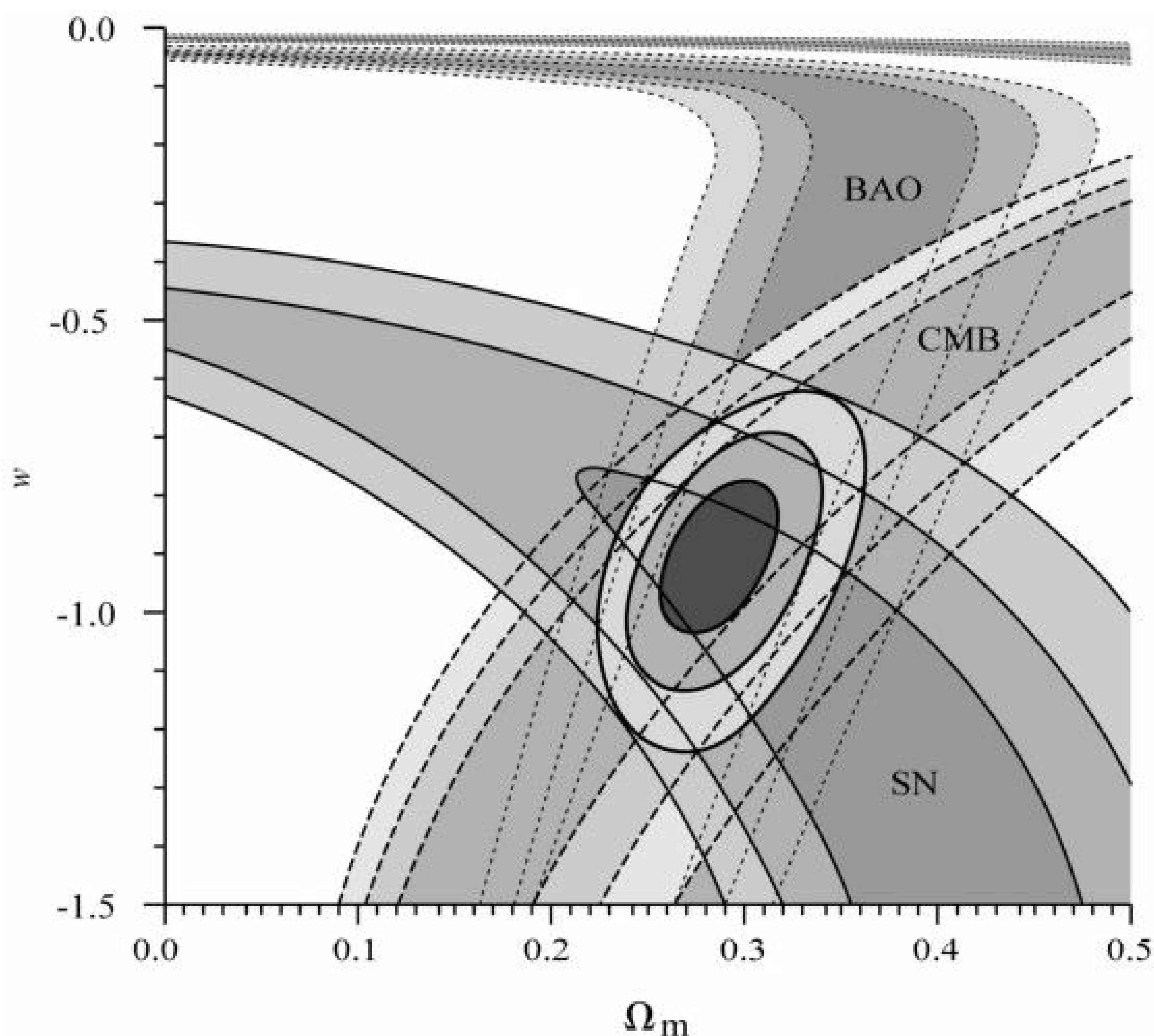


图12.6 超新星观测数据（SN）、微波背景辐射的涨落（CMB）和重子声学振荡效应（BAO）对暗能量的压强与能量密度的比值w以及其他物质占宇宙总能量的比例Ω_m所施加的限制

加速膨胀的宇宙将会迎来什么样的命运？这是其中最引人注目的一个问题。1986年，弗兰克·梯普勒和我证明，在未来某一天，无论是何

种形式的计算机或“大脑”，任何信息的处理过程最终都会停下来。恒星和星系不再形成。宇宙膨胀得太快了，以至于任何物理过程都无法持续进行。宇宙的未来看起来像是一个无穷无尽的公墓，里面是死去的恒星和孤立的基本粒子，而且它们一定会永远地“与世隔绝”。加速膨胀就意味着任何观测者眼前都会出现一个视界，在视界之外，什么都看不到。这种情形就像身处一个大型黑洞的内部。如果宇宙的空间在加速膨胀，从很远的地方发出来的光就永远无法到达我们这里。这种奇怪的现象最先是由阿瑟·爱丁顿在1933年发现的。在他的名著《膨胀的宇宙》中，他这样描述这种现象：

星系和星系之间的距离最终将变得如此之大，相互之间的退行运动将变得如此之快，以至于光或者其他因果效应都来不及从一个星系传播到另一个星系中。星系之间的所有联系都中断了；每个星系都是一个独立王国，不受外界的任何影响。这样的崩溃是一场令人无法想象的噩梦，尽管这并不会给人类的命运带来任何具体的灾难。[13]

如果你担心多重宇宙会在我们的宇宙图景中引入无法观测的部分，那么你也得担心，即使在“一次性”的宇宙中，现在或无穷远的将来都会存在我们无法看到的角落。

然而，尽管我们的观测结果和爱因斯坦的一个简单模型符合得很好，其中仍然有很多未解的难题。在过去的50亿年中，有两种形式的能量影响了宇宙的膨胀速率。一旦恒定的真空能控制了局面，它的影响就会越来越强大，而所有形式的物质和辐射都会随着宇宙的膨胀被迅速稀释。梯普勒和我证明了，由此引发的结果非常重要。[14]一旦宇宙开始加速膨胀，由牛顿、金斯和栗弗席兹发现的引力不稳定过程中的不规则性就会停止增长。不规则性的增长跟不上膨胀速率的增加。如果到了这个大变革时代仍然没有形成星系，那就永远也形成不了。没有星系，就不会形成稳定的恒星，也不会产生碳元素以及创造其他生命和观测者所需的砖块。

迷雾重重的宇宙

……宇宙学常数一直以来就像你祖父的一对鞋套，偶尔会因为衣着搭配的需要而拿出来穿一穿。但最新的结果表明，它不仅重新流行了起来，更是成了时尚的必备元素。

——罗伯特·克尔什纳

宇宙学常数不但存在，而且还在宇宙中占主导地位，这一发现震惊了许多天文学家和大部分粒子物理学家。如果以宇宙学常数可能取到的最大值为单位，就可以将它表示成一个无量纲量。这个自然常数表示宇宙的真空能有多小，取值范围为0到1。

直到第一个直接证据，也就是超新星的观测结果出现以前，关于微波背景辐射和星系结团的不同类型观测的最佳拟合结果也暗示了宇宙学常数的存在，但这并不能让粒子物理学家信服。粒子物理学理论预示了宇宙学常数的存在，但其值在0到1之间且预期等于1。然而，在1998年超新星数据之前的观测数据所允许的宇宙学常数最大不超过10^{-120}，这就

跟0差不多了。因此，大多数粒子物理学家相信，宇宙学常数就是0，其原因一定是高能物理学中存在某种尚未发现的重要原理。在得知超新星观测证实了宇宙中的确存在宇宙学常数，或者某种与之类似的机制，而且其取值离奇小，只有 10^{-120}时，人们不禁百思不得其解。为什么这么小？什么机制能够产生这么小的宇宙学常数，比粒子物理学家预期的取值还要小10^{120}倍？

宇宙学常数带来的重重迷雾还不止于此。如果这个微小的数目再扩大10倍，即不过10^{-119}时，星系和恒星就无法形成了。而且，即使你举手回答说宇宙学常数一开始的时候就是10^{-120}，这也无法解决问题。宇宙的甚早期存在一系列特殊时期，那时，自然界的基本相互作用逐个从一种统一的相互作用中剥离了出来，于是不同相互作用的强度也出现了差异。每当这个时候，宇宙学常数的大小就会被重置，变得比初始的10^{-120}大得多。任何物理机制，如果想要解释宇宙学常数今天的大小，都必须把眼光放长远，横跨宇宙的不同历史时期，提前预知宇宙不同时期的不同变化结果，并将这些影响全部抵消。我们从没听说过有这样一种物理机制的存在。

到目前为止，没有任何人能够解释，为什么宇宙学常数或者我们称之为“暗能量”的取值如此怪异。也许引力理论中存在一些完备的新机制，或者量子的机制，能够解释这种现象。物理学家开始尝试拓展爱因斯坦的引力理论，向其中加入新型的几何形式。这样的话，物质所引起的空间弯曲程度就会略有不同，当宇宙的体积变大时，这种差异也会随之增大。[15]人们估计，如果考虑到量子宇宙学和初始奇点极端条件之间的相互影响，引力理论就会发生变化。然而，人们并不指望爱因斯坦理论的变化发生在宇宙非常晚的时期，因为此时物质的密度很低，产生的引力场也很弱，没有什么量子引力的效应值得考虑。这样的理论修正当然会导致宇宙晚期的加速膨胀[16]，不过其中大部分修正也会在别的地方引起一些我们还没看到的效应。然而，经过这般修正的理论看起来就跟宇宙学常数一样不自然。因此，暗能量是现代宇宙学的前沿问题。

到目前为止，弦理论和M-理论的研究还没能对这个问题提供什么帮助，这令人很失望。也许，我是说也许，我们走错了方向，对于宇宙的加速膨胀或宇宙学常数来说，并不存在传统意义上的解释。宇宙学常数的大小也许是随机的，是由多重宇宙弦景观的某种量子机制所导致的。在不同的宇宙中，宇宙学常数的大小是随机变化的。我们也许会发现自己生活在多重宇宙的一个子宇宙中，其中宇宙学常数刚好允许星系和恒星的形成。但是对于大多数研究者来说，这个结果非常令人失望，可是我们得学着适应。如果有人因此畏缩不前，认为这是科学方法的溃败，那我会给他讲一个小故事，一个科学研究的真实故事。

想象你在公元1600年，正试图说服开普勒（1571～1630），他的太阳系理论并不需要预言太阳系里有多少行星。但在开普勒看来[17]，行星的数量表明自然界中存在一种深层的数学对称性。如果谁认为太阳系的一个数学模型并不能完全且唯一地确定行星数目的话，那他就是大逆不道，就是在否定科学。

开普勒发现，五种正多面体都有一组唯一的内嵌球或外接球。如果这些圆球和正多面体像洋葱皮一样一个接一个地嵌套起来，一共就会形成六层圆球，对应当时已知的六颗行星——水星、金星、地球、火星、木星和土星。按照正确的顺序摆放正八面体、正二十面体、正十二面体、正四面体和立方体，开普勒发现，这些外接球面之间的间距就跟每个行星（假设的）圆形轨道半径的比例一致。他将这个现象画成了图12.7。

今天，任何头脑正常的行星天文学家都不会像开普勒那样，试图预测太阳系中有多少颗行星，或者寻找某种深层意义。那个数目不过是历史的巧合，源自于太阳系早期物质的合并和另外一些偶然因素。其实，太阳系完全可以是另外一个样子。行星的数目应该是随机的。所以行星天文学家真正关心的问题是，随着轨道半径的不同，行星的大小和成分会体现出什么样的变化趋势以及其他的动力学细节。

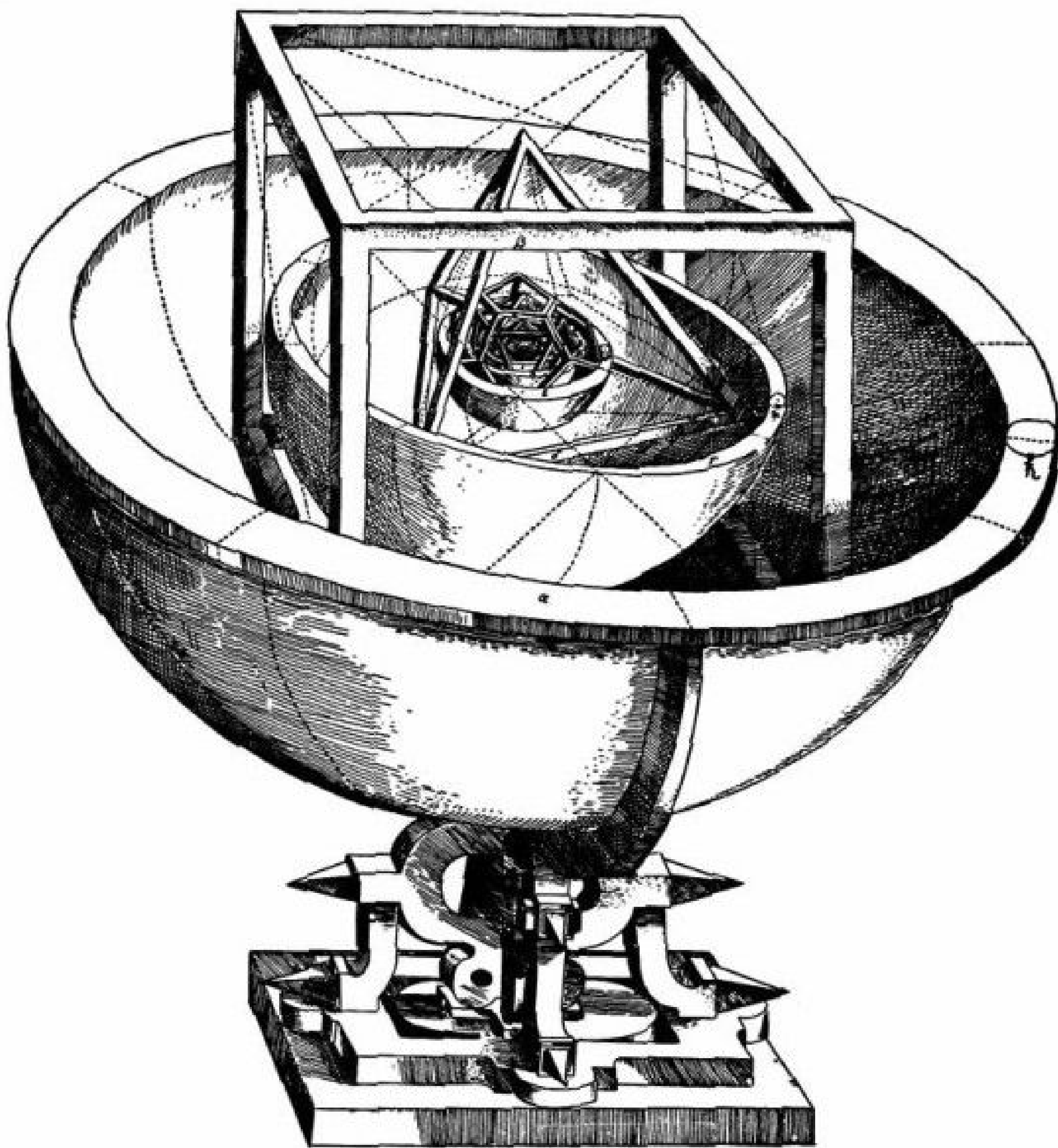


图12.7 开普勒对柏拉图太阳系模型的扩展，摘自他第一部重要的天文学著作《宇宙的奥秘》（Mysterium Cosmographicum，1600）

宇宙学常数会不会像行星总量一样呢？仅仅由于我们认为宇宙学常数是宇宙的一个基本性质，并不能说明它的取值必然是唯一的，必然由自然法则完全确定。宇宙学常数可能是那些深层规律所产生的一个完全随机的结果：一个意外，尽管这个意外对我们非常重要。除此之外，并无更多意涵。但就像太阳系的故事一样，这并不是整个故事的结局。可

见宇宙还存在很多特征可以作为科学研究的实验室，要想知道我们的理论和解释对不对，就必须把这些特征都找出来。或许在遥远的将来，宇宙学家看待我们对宇宙学常数的研究，就像我们看待开普勒专注于太阳系的行星数量一样。

但或许还有一线希望。最近，道格拉斯·肖和我为这个问题[18]提出了一个全新的解决方案。我们向爱因斯坦提出的方程组中加入了一个方程，于是就允许宇宙学常数在必要时改变它的能量形态，并且会保证这种非量子的图景能够决定宇宙近期的膨胀行为。由此，我们可以得出一个漂亮的预言。如果你是宇宙诞生后 t_u时刻的观测者，你总是可以观测到宇宙学常数的效应，其大小为（t_p/t_u）^{2}，其中$t_p=10^{-43}$秒叫做普朗克时间。由于目前宇宙的年龄是$t_u=4.3\times10^{17}$秒，因此我们预期宇宙学常数的大小为0.5×10^{-121}，与实际观测相符。[19]值得注意的是，我们并没有引入任何新的能量形式，也没有对爱因斯坦方程组作出量子修正，也没有涉及所有可能宇宙形成的多重宇宙，没有让人择原理要求其中的宇宙学常数取值很小。更重要的是，这个新理论对可见宇宙空间曲率的当前取值作出了一个明确的预言：曲率应该是正的，而且曲率项对应的等效能量的观测值[20]应该是-0.005 6。该能量现在的观测值在-0.013 3到+0.008 4之间。不过，综合欧洲空间局的普朗克卫星绘制的全天空微波背景辐射分布图以及图12.5中的其他观测数据，将会大幅增加观测的精度。我们应该很快就能证实或证伪这个十分精确的预言了。

我们的故事讲了很久。开始时，我们讲的是要想理解天上的事情是多么困难。我们又讲了确认恒星和星系是否存在是多么困难的一件事情。1915年，爱因斯坦广义相对论的创立是一个分水岭，标志着我们对宇宙以及对其他可能的宇宙有了全新的看法。那是我们第一次开始从整体上研究宇宙，我们作出了各式各样的理论预言，理解了不同宇宙模型的主要性质，并有所取舍。我们也讲到，爱因斯坦方程组有越来越多的解被揭开了面纱，那些宇宙模型向人们展示了大量出人意料的特性。我们也讲到，人们如何用方程组的解来解释天文学观测，或者用哲学观点来描述宇宙应该是什么样子。

现有的宇宙学标准模型是我们说过的最长寿的模型，也就是暴胀宇宙理论。这种理论认为，宇宙在很久很久以前曾经爆发过一次加速膨胀，并且它能成功地预言我们观测到的宇宙早期产生的微波背景辐射的微小涨落服从某种特殊的模式。用超级计算机模拟星系结团的复杂行为，再将结果与新型望远镜无与伦比的精确观测相比较，我们就绘制出一幅宇宙的图景，有些地方我们都能理解，有些地方仍然迷雾重重。

我们不得不面对这样一种观点，即多重宇宙中有无穷多个真实的宇宙，每个宇宙的性质都各不相同，我们的宇宙只是其中之一。我们的宇

宙某些方面应该很特别，这才会导致我们的存在，或者还有其他形式智慧生物的存在。如今我们又发现，大约在50亿年以前，我们的宇宙开始了第二轮加速膨胀。

哥白尼告诉我们，我们的行星并不是宇宙的中心。现在，我们也许不得不承认，就连我们的宇宙也不是宇宙的中心。

注释

[1] 转引自：C. Brownlee, ‘Hubble’s Guide to the Expanding Universe’, National Academy of Sciences Classics, online at http://www.pnas.org/misc/classics2.shtml.

[2] C. Green,*The Human Evasion, H*amish Hamilton, London (1969), p. 12.

[3] 会议论文集可见：*Critical Dialogues in Cosmology,* ed. N. Turok, World Scientific, Singapore (1997).

[4] M. Davis, G. Efstathiou, C. Frenk and S.D.M. White, *Astrophys. J.*292, 371 (1985).

[5] A. Riess et al.,*Astron. J.*, 116, 1009 (1998), and S. Perlmutter et al.,*Astrophys. J.* 517, 565(1999).

[6] 将这种超新星看作标准烛光的提议是大卫·阿奈特（David Arnet）最早在 1979年提出的，不过一直以来人们认为实现起来难度太大。

[7] 在量子力学中，如果试图将不同的电子放入同一个状态，就会产生这种抵抗性压强，叫做电子简并压强。在由此而产生的白矮星中，原子和电子都尽可能地靠在一起。

[8] 如果宇宙是平坦的话，目前对这组数据最精确的估计是，暗能量占72.1%±1.5%，其他物质占27.9%±1.5%，置信度为68%。

[9] 这是压强和能量密度之间可能存在的最简单的关系。也有人研究了更加复杂的关系。

[10] G. Lemaître,*Proc. Natl. Acad.Sci.* 20, 12 (1934).

[11]方程$a(t) = a_0\left[\sinh\left(t\sqrt{3\Lambda}/2\right)\right]^{2/3}$描述了图3.13中的曲线，其中$a_0$是宇宙现在的半径， a(t)是t时刻宇宙的半径，t是宇宙自膨胀以来所花的时间。

[12]这个巧合是爱丁顿最先发现的，他在一篇关于爱因斯坦静态宇宙不稳定性的论文［*Mon. Not. Roy. Astron. Soc.* 90, 677 (1930)］中提到了这一发现。他说，我们应该庆幸自己是“多么幸运”，赶上了宇宙美景，“及时地观察到了天空有趣而日渐暗淡的特征”。

[13] A. S. Eddington,*The Expanding Universe*, Cambridge University

Press, Cambridge (1933), pp. 85–86.爱丁顿将加速宇宙相互远离的部分叫做“泡泡”，并将它们定义为“不存在因果联系的区域”。在那本书那一章的题头中，爱丁顿引用了弗朗西斯·培根一首著名的诗《人的一生》（‘The Life of Man’）：“世界是个泡泡，人生/不过一拃。”“泡泡”这个说法似乎典出于此。

[14] J. D. Barrow and F. J. Tipler, *The Anthropic Cosmological Principle*, Oxford University Press, Oxford (1986), section 6.9.

[15] S. M. Carroll, V. Duvvuri, M. Trodden, and M. S. Turner,*Phys. Rev*. D 70, 043528 (2004).

[16]有一种很简单的情况能说明，在爱因斯坦理论的某种推广理论之中，仍然存在德希特加速膨胀的宇宙解。参见：J. D. Barrow and A. Ottewill, *J. Phys.* A 16, 2757 (1983).事实上，能够解释加速膨胀的爱因斯坦理论的大部分推广理论都可以等效地看作是往爱因斯坦理论中加入了某种形式的物质，证明过程参见：J. D. Barrow and S. Cotsakis, *Phys. Lett*. B 214, 515 (1988).有意思的是，这种物质和宇宙早期历史中产生暴胀的物质是一类的。

[17] M. Caspar,*Kepler*, transl. C. D. Hellman, Dover, New York (1993); O. Gingerich,*The Eye of Heaven: Ptolemy, Copernicus, Kepler,*Springer, New York (1997).

[18] J. D. Barrow and D. J. Shaw, ‘A New Solution of the Cosmological Constant Problems’ (2010), arXiv:1007.3086; D. J. Shaw and J. D. Barrow, ‘A Testable Solution of the Cosmological Constant and Coincidence Problems’ (2010), arXiv:1010.4262.

[19]宇宙当前的年龄之所以会在这个预言中出现，是因为唯一对宇宙学常数取值有贡献的是那些与我们有因果联系的作用源，这个范围大致等于ct_u。

[20]这个量就是空间曲率当前的引力效应对应的能量密度占宇宙总能量密度的比例，其表达式为$\Omega_k=-k/a^{22}H$（其中的变量为取作当前的大小）。这类似于图12.5中画的物质和宇宙学常数对宇宙总能量密度的贡献。注意，预言中的$\Omega_k=-0.005\ 6$是负数，所以k必须大于零。

图片版权

Fig. 1.1 Reproduced from J.D. Barrow, *The Artful Universe*, Clarendon Press, Oxford (1995).●**Fig.1.2** Reproduced from Archie E.Roy, 'The Origin of the Constellations', *Vistas in Astronomy*, 27, 171–197. Copyright © 1984, reprinted by permission of Elsevier. ● **Fig.1.4** Herman Eisenbeiss/Science Photo Library.●**Fig. 1.6a** Henry E. Huntington Library and Art Gallery.●**Fig. 1.9** Royal Astronomical Society/Science Photo Library; from Nicolaus Copernicus, *De revolutionibus orbium coelestium* (1543). ● **Fig. 1.10** Reproduced from E. R. Harrison, *Cosmology: The Science of the Universe*, Cambridge University Press, Cambridge (1981).●**Fig.2.2** Reproduced from E. R. Harrison, *Cosmology: The Science of the Universe*, Cambridge University Press, Cambridge (1981).●**Fig. 2.3** akg-images.●**Fig. 2.4** Reproduced from E. R. Harrison, *Cosmology: The Science of the Universe*, Cambridge University Press, Cambridge (1981). ● **Fig. 2.5** Reproduced from A. R. Wallace,*Man's Place in the Universe*, Chapman and Hall, London (1912). ● Fig. 2.7b Lynne McDonagh.●**Fig. 2.9** Archive, Astrophysikalisches Institut Potsdam (AIP).●**Fig 3.2** Courtesy of the Albert Einstein Archives, The Hebrew University of Jerusalem, Israel. ● **Fig. 3.9** akg-images. ● **Fig.3.12** Reproduced from A.V. Douglas,*The Life of Arthur Stanley Eddington*, Nelson, London (1956).●**Fig.3.14** Science Photo Library. ● **Fig.3.15** Leiden Observatory Archives; New York Times Photo Archives/REDUX/Eyevine. ● **Fig. 4.1** Mary Lea Shane Archives/Lick Observatory. ●**Fig. 4.2** Reproduced from E.R. Harrison, *Cosmology: The Science of the Universe, Cambridge University* Press, Cambridge (1981). ● **Fig. 4.3** Pasieka/Science Photo Library. ● **Fig. 4.4** Courtesy of University Archives, Columbia University in the City of New York.● **Fig.4.6** Roger Viollet/Getty Images. ● **Fig.4.7** Markus Pössel (MPIA and AEI)/http://www.einstein-online.info. ● **Fig. 5.1** Courtesy of Professor Daniel Straus.●**Fig. 5.2b** Charles Dyer and Allen Attard.●**Fig. 5.4** John D. Barrow.●**Fig. 5.6** Time & Life Pictures/Getty Images. ● Fig. 6.1 Reproduced by permission of the Master and Fellows of St John's College, Cambridge.●**Fig. 6.2** American Astronomical Society/NASA Astrophysics Data System. Reproduced from *Astrophysical Journal*, vol. 94, no. 3, 1941. ● **Fig. 6.3** American Astronomical Society/NASA Astrophysics Data System.

Reproduced from *Astrophysical Journal*, vol. 94, no. 3, 1941. ● **Fig. 6.5** Reproduced courtesy of Dr Arno Penzias.●**Fig. 6.6** Cover of R. A. Alpher'sGenesis of the Big Bang, Oxford University Press, Oxford (2001); cover design by Emily Kulp.● **Fig. 6.7** Science Photo Library. Reproduced from R.V. Wagoner, W.A. Fowler and F. Hoyle, 'On the Synthesis of Elements at Very High Temperatures', *Astrophysical Journal*, vol. 148, no. 3, 1967. ● **Fig. 7.1** AirTeamImages, copyright © Steve Morris 2010.●**Fig. 9.9** Reproduced from M. Lachièze-Rey, J.-P. Luminet and J. Laredo,*Celestial Territory: From the Music of the Spheres to the Conquest of Space*, Cambridge University Press, Cambridge (2001)●**Fig. 9.10** Copyright © ARS, NY and DACS, London 2010; image copyright © The Metropolitan Museum of Art.●**Fig. 11.1** Mehau Kulyk/Science Photo Library.●**Fig. 11.3** Emilio Segrè Visual Archives/American Institute of Physics/Science Photo Library.●**Fig. 11.4** Reproduced courtesy of Kip S. Thorne.

译后记

非洲是人类的发源地，但人类走出了非洲。“地球是人类的摇篮，但人类不可能永远生活在摇篮中。”（航天之父齐奥尔科夫斯基的名言。）所以即使从功利的角度考量，我们也要关心宇宙，关注宇宙学的发展。

我是一名宇宙学爱好者。但长期以来我只知道宇宙学的某些结论，不知道这些结论是怎么来的，更没有大局观。就像一个从小衣来伸手饭来张口的孩子，知道什么好吃，却不知道好吃的东西“粒粒皆辛苦”，不知道我们的祖先“暴霜露，斩荆棘，以有尺寸之地”。

关于宇宙最精确的描述叫做“标准宇宙学模型”。当我在 2010年年底第一次看到这本书时，还以为它讲的就是这个标准的模型。完成了大部分翻译工作后，我才意识到本书主要讲的是非标准的宇宙学模型。就像三国演义之中英雄辈出，作者不会因为司马氏统一了天下，就把魏蜀吴的兴衰一笔带过。

宇宙学的科学理论也像历史脉络一样匪夷所思。例如，为什么科学家对大爆炸理论的评价很高？这并不符合直觉，几乎每个人刚听到这个理论时都不愿相信。甚至还有很多热心读者提出了自己的见解，要“推翻”大爆炸理论，将宇宙学“引入正确的轨道”。读过这本书以后，我们就会了解，大爆炸理论就像《三国演义》中的司马懿，一开场时并不是主角。他有很多敌人，敌人不会自动缴械投降，而且新的敌人会不断出现。但最终他把那个年代的敌人都打败了。他的地位不是无缘无故来的，而且他的手下败将都是当时数一数二的人物。

总有一天，大爆炸理论肯定也会被取代。不过，取代它的理论同时一定又是它的忠实继承者。本书的作者专门用了一节的内容向大家解释，新理论不会推翻旧理论。新理论只不过向人们指出了旧理论的局限性，并将旧理论的正确预言作为自己的一个特殊情况。

随着我们对宇宙的了解日益丰富，我们发现拦在路上的未解难题越来越多。每一天都有科学家撰写论文，试图解决其中某个角落的问题。一个个比科幻小说更离奇的宇宙模型不断在学术期刊上出现。其中大部分的研究将来都会是“浪淘尽，千古风流人物”，胜出的总是少数。所以，这仍然是一个群雄逐鹿的时代。欢迎大家加入这场智力的角逐！

译者对语言的把握远逊于对物理的把握，尽管绞尽脑汁完成了这部书的翻译，又反复校对了几遍，但仍然会存在有待改进的地方。还请发现了问题的读者不吝赐教。在这里要特别感谢我的朋友 Yao Zhang、

popmary、yuler、王道还，科学松鼠会的红猪、Steed、Ent、老孙、小菊、odette、0.618，网友冰激凌太妹、观雏鸡可以识仁、Simon Bridge，以及宇宙学专家王一、蔡一夫和张旭给予的帮助，他们的建议使此译本增色不少。还要感谢本书的编辑楼伟珊，他的专业素养和敬业精神使我受益良多。

李剑龙

2012-2-8